



Paulo Jorge Silva Póvoa

Nº 41977

Extração automática de elementos em fotografia aérea por fotogrametria

Dissertação para obtenção do Grau de Mestre em
Engenharia Informática

Orientador : Prof. Doutor Fernando Pedro Reino Silva Birra,
Prof. Auxiliar, DI/FCT/UNL

Co-orientador : Tenente-Coronel de Artilharia Engenheiro Geógrafo,
Rui Alberto Coelho Dias, Instituto Geográfico do
Exército

Júri:

Presidente: Prof. Doutora Maria Cecília Gomes

Arguente: Prof. Doutora Paula Redweik

Vogal: Prof. Doutor Fernando Birra



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Dezembro, 2014

Extração automática de elementos em fotografia aérea por fotogrametria

Copyright © Paulo Jorge Silva Póvoa, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

À minha filha Alice

Agradecimentos

Gostaria de iniciar os agradecimentos pela cooperação que me foi prestada pelo Instituto Geográfico do Exército, sobretudo pelo auxílio prestado pelos seus colaboradores. Nomeadamente na pessoa do Senhor Tenente-Coronel Rui Dias, por ter aceitado ser co-orientador desta dissertação e pelos esclarecimentos prestados em termos de fotogrametria. Não quero deixar de agradecer, em particular, aos colaboradores da secção de fotogrametria, por terem dispensado do seu tempo de trabalho e pessoal para contribuir nesta dissertação.

À professora Doutora Paula Redweik, por me ter auxiliado em momentos em que estava "quase tudo perdido", foram duas meias horas que valeram por muito, pois permitiram compreender conceitos, que para mim era um completo mistério.

Aos meus colegas de faculdade, com quem tive a oportunidade de conviver durante estes anos, de dia e noite.

Aos meus pais, pela educação e valores que me transmitiram, pelas discussões e carinho que me deram e pela paciência que tiveram para comigo.

À minha esposa, por todo o apoio que me foi dado e pelas responsabilidades que teve de assumir na minha ausência. Por todas as críticas e elogios que fizeram com que este trabalho fosse realizado.

À minha filha Alice, que teve um papel fundamental para manter a minha sanidade, sempre que me pedia para ir brincar com ela.

Resumo

Na cadeia de produção de cartografia digital, que se encontra implementada no IGeoE, existe a necessidade de efetuar a restituição de objetos através da vetorização 3D. Sendo este processo completamente manual, o operador identifica e adquire os objetos, em ambiente estereoscópico, segundo as regras estabelecidas nas normas de aquisição.

A aquisição de construções, nomeadamente os edifícios, são dos objetos que mais recursos consomem devido à sua frequência e dificuldade de restituição. A possibilidade de simplificar esta parte do processo proporciona um benefício substancial para toda a cadeia de produção.

Pretende-se assim detetar edifícios em fotografias aéreas, extraíndo a sua informação planimétrica e altimétrica, para posterior inserção num SIG. Para a obtenção da altimetria são utilizados os princípios de fotogrametria analítica através das equações de colinearidade.

Este problema torna-se relevante devido ao facto de se pretender retirar informação de fotografias, que possuem bastante informação, com recurso à computação gráfica, através de técnicas de segmentação em vários níveis e fotogrametria, juntando assim duas áreas do saber.

Esta solução permite automatizar um processo que é predominantemente manual, contribuindo para melhorar a cadeia de produção sem a alteração de funcionamento da mesma.

Palavras-chave:

Edifícios, Segmentação, Fotografia aérea, Fotogrametria, Equações de colinearidade, SIG.

Abstract

In the production process of digital mapping that is implemented in IGeoE, there is a need to form the restitution of objects through 3D vectoring, this process is completely human craft . The operator identifies and acquires objects, in stereoscopic environment according to the rules of acquisition.

The construction acquisition, including buildings, are objects that consume most of the resources due to their frequency and difficulty of restituting the simplification of this part of the process provides a substantial benefit for the entire production process.

The aim is to detect buildings structures in aerial photographs , extracting the planimetric and altimetry information of the structure for later insert into in a GIS . For obtaining the altimetry will be used principles of analytical photogrammetry through the collinearity equations .

This problem becomes relevant due to the fact if you want to retrieve information from photographs , we have enough information , using the computer graphics , through techniques targeting at various levels and photogrammetry , thus joining two areas of knowledge .

This solution allows the possibility to automate a process that is predominantly human craft , contributing to improve the production process without changing it.

Keywords: Buildings, Segmentation, Aerial Photography, Photogrammetry, Collinearity equations, GIS.

Conteúdo

Abreviaturas	xvii
Glossário	xix
1 Introdução	1
1.1 Enquadramento	1
1.2 Motivação	4
1.3 Objetivos	4
1.4 Contribuições esperadas	5
1.5 Organização do documento	5
2 Restituição de objetos	7
2.1 Câmaras fotográficas	7
2.2 Operações preliminares	9
2.2.1 Aerotriangulação	13
2.3 Restituição	16
2.3.1 Normas	17
2.3.2 Equipamentos	17
2.3.3 Processo	18
2.4 Equações de colinearidade	21
3 Trabalho Relacionado	27
3.1 Extração com base em características	28
3.2 Extração baseada em métodos probabilísticos	33
3.3 Sombras	34
3.4 Organização de cores	37
3.5 Detecção de pontos caraterísticos	40
3.5.1 <i>Scale-invariant feature transform</i>	41
3.5.2 <i>Speeded Up Robust Features</i>	41

3.6	Correlação de pontos	42
4	Trabalho Realizado	45
4.1	Segmentação da fotografia	47
4.1.1	MeanShift	48
4.1.2	Quantização do espaço de cor	50
4.1.3	Segmentação de arestas	52
4.1.4	Mapa de disparidades	56
4.2	Construção de polígonos	57
4.2.1	Deteção de contornos	58
4.2.2	Deteção de arestas	60
4.2.3	Deteção de vértices	62
4.2.4	Ordenação de vértices	63
4.2.5	Deteção de contornos (Convex hull)	65
4.3	Determinação de coordenadas terreno	65
4.3.1	Transformações afins	66
4.3.2	Identificação de pontos homólogos	68
4.3.3	Coeficiente de correlação	71
4.3.4	Conjugação de dois métodos	73
4.3.5	Equações de colinearidade	75
4.3.6	Outros processos	76
5	Avaliação	81
5.1	Áreas avaliadas	81
5.1.1	Zona residencial	82
5.1.2	Zona urbana	84
5.1.3	Zona histórica	84
5.1.4	Zona rural	84
5.2	Taxa de identificação	84
5.2.1	Zona residencial	85
5.2.2	Zona urbana	87
5.2.3	Zona histórica	88
5.2.4	Zona rural	90
5.3	Análise pelas fases do método	91
5.4	Taxa de aceitação	96
6	Conclusões e trabalho futuro	101
6.1	Conclusões	101
6.2	Trabalho futuro	102
6.2.1	Aspetos gerais	102
6.2.2	Segmentação da imagem	103
6.2.3	Construção de polígonos	104

6.2.4	Determinação de coordenadas terreno	104
A	Algoritmo probabilístico da transformada de Hough	107
B	Análise equações de colineariedade	111
C	Imagens da análise por fases	113
D	Imagens da taxa de aceitação	119

Abreviaturas

- CAD - *Computer-aided Design*.
- CIE - *Commission Internationale de l'Eclairage*.
- DAD - Departamento de Aquisição de Dados.
- DCT - *Discrete Cosine Transform*.
- DGT - Direção Geral do Território.
- DoG- *Difference-of-Gaussian*.
- GIS - *Geographic Information System*
- GNSS - *Global Navigation Satellite System*.
- GPU - *Graphics Processing Unit*.
- (GSD) - *round sample distance*.
- EMQ - Erro Quadrático Médio.
- FLANN - *Fast Library for Approximate Nearest Neighbors*.
- IGeoE - Instituto Geográfico do Exército.
- ISAE - *ImageStation Automatic Elevation*.
- ISAT - *Image Station Automatic Triangulation*.
- MDS - Modelo digital de superfície.
- MDT - Modelo digital de terreno.
- MMQ - Método dos Mínimos Quadrados
- NATO - *North Atlantic Treaty Organization*.

- nDSM - *Normalized Digital Surface Model*.
- NDVI - *Normalized Difference Vegetation Index*.
- OpenCV - *Open source Computer Vision*.
- pBased - Modelo de cores apresentado no artigo [\[CGZ08\]](#)
- PP - Ponto Principal.
- QGIS - Quantum GIS
- SERVIR - Sistema de Estações de Referência GNSS VIRTuais
- SIFT - *Scale-invariant feature transform*.
- SFot - Secção de Fotogrametria.
- SGM - *Semi-Global Matching*.
- SIG - Sistema de Informação Geográfica.
- STANAG - Standardization Agreement.
- SURF - *Speeded Up Robust Features*.
- SVM - *Support Vector Machine*.
- TIN - *Triangulated Irregular Network*.
- WGS - *World Geodetic System*.

Glossário

- **Azimute:** Ângulo entre a direção de referência e a direção pretendida, no sentido horário.
- **Coordenadas imagem(fotocoordenadas):** Localização de um ponto num sistema de retangular bidimensional, que representa o plano fotográfico.
- **Coordenadas objeto ou terreno:** Localização de um ponto num sistema retangular tridimensional, que define a localização do ponto no espaço objeto.
- **Espaço/sistema/espectro de cores :** Organização específica de cores, num referencial, na maior parte dos casos, tridimensional.
- **Estereorrestituição:** Restituição de elementos em ambiente estereoscópico. No IGeoE é efetuada com recurso a fotogrametria digital.
- **Fotografia origem:** Designação da fotografia utilizada para a obtenção de informação 2D.
- **Fotografia destino:** Designação da fotografia que permite a determinação da informação 3D, conjuntamente com a informação da fotografia origem.
- **Imagem binária:** Imagem digital onde os píxeis podem somente ter dois valores, branco ou preto.
- **Imagens em pirâmide:** Técnica que permite armazenar a informação de uma fotografia, por níveis, cada um com uma dimensão diferente.
- **Informação geoespacial:** Informação sobre a posição relativa de objetos na superfície terrestre.
- **ISAE:** Programa utilizado no IGeoE para efetuar os modelos digitais do terreno.
- **ISAT:** Programa utilizado no IGeoE para efetuar o processo de aerotriangulação.

- **GNSS** : Técnica de posicionamento de navegação que permite a utilização de várias constelações de satélites.
- **Ground Sample Distance**: Distância entre o centro de píxeis adjacentes medida no terreno.
- **Matriz Fundamental**: Matriz 3×3 que relaciona os pontos homólogos num conjunto fotografias *stereo*. Com esta matriz e as coordenadas homogêneas, de um ponto, determina-se a linha epipolar, na segunda fotografia, onde o ponto homólogo se encontra.
- **Matriz Essencial**: Matriz 3×3 que relaciona os pontos homólogos num conjunto de fotografias *stereo*, assumindo o modelo *pinhole* de câmara e que esta se encontra calibrada. Esta é uma especialização da matriz fundamental para o processo de coordenadas normalizadas da imagem. Pode ser utilizada para determinar a posição relativa e a orientação de duas câmaras.
- **Modelo estereoscópico** ou **estereomodelo**: Conjugação de duas fotografias sucessivas e com sobreposição, que permitem visualizar uma imagem tridimensional
- **Modelo Digital de Superfície**: Modelo que descreve a superfície, incluindo todas as estruturas presentes acima do terreno.
- **Modelo Digital de Terreno**: Modelo que descreve o terreno, não incluindo nenhuma estrutura presente acima do terreno.
- **nDSM**: Modelo de superfície que contém a diferença entre o MDS e MDT;
- **NDVI**: Indicador gráfico com a informação da vegetação.
- **Paralaxe**: Ângulo formado por dois raios projetivos dos centros óticos, das fotografias, para o mesmo ponto no terreno.
- **Polígono complexo**: Um polígono que não seja convexo ou côncavo designa-se por polígono complexo, em computação gráfica. Estes incluem polígonos que se auto intersetem ou possuam "buracos" no seu interior.
- **Ponto imagem**: Ponto no sistema de coordenadas imagem relativo a um sistema retangular bidimensional.
- **Ponto objeto**: Ponto no sistema de coordenadas objeto relativo a um sistema retangular tridimensional.
- **Pontos caraterísticos**: Pontos de uma fotografia, que possuem um conjunto de características que podem ser utilizadas para efetuar a sua classificação. Estes podem ser denominados ainda por *keypoints*, pontos chave, pontos de interesse, identificadores em outros trabalhos.

- **Pontos homólogos:** Pontos, em imagens diferentes, que são semelhantes ou têm afinidades.
- **Projeto SERVIR:** Rede de estações de referência GNSS instaladas em pontos conhecidos, que permite a correção de erros no sinal recebido dos sistemas de posicionamento.
- **Restituição:** Processo de extração de informação de um modelo estereoscópico para informação vetorial.
- **SIG3D:** Projeto que decorre no IGeoE para migrar o processo de aquisição de dados de CAD para SIG, por processos fotogramétricos.
- **Vetorização 3D:** Transformação para suporte digital dos elementos relevantes para a carta militar 1:25 000, em 3 dimensões.
- **WGS84:** Sistema de coordenadas projetadas, que utiliza um elipsoide de referência de origem geocêntrica.
- **WGS84 militar:** Sistema de coordenadas WGS84 com uma translação da origem das coordenadas, que é utilizado no IGeoE. Este utiliza o Datum WGS84 com projeção transversa de Mercator com uma falsa origem, para permitir que as coordenadas em Portugal sejam positivas.



Introdução

Este relatório insere-se na dissertação do Mestrado Integrado em Engenharia Informática, a qual visa desenvolver um método que permita a vetorização 3D de objetos, presentes em fotografias aéreas, para a Carta Militar de escala 1:25 000.

Esta dissertação é de âmbito empresarial e encontra-se integrada no Instituto Geográfico do Exército (IGeoE).

Para a produção de cartografia, o IGeoE possui um catálogo de objetos que define os elementos a representar e a forma como são restituídos. Estes podem ser restituídos como pontos, linhas ou polígonos.

No caso particular desta dissertação estamos interessados em efetuar a restituição de polígonos. Mais concretamente, polígonos que correspondem a edifícios, que constam da informação da Carta Militar. Pretende-se ainda obter a informação altimétrica dos mesmos, mantendo a consistência dos dados produzidos atualmente.

O método deve possuir um conjunto de mecanismos automáticos, minimizando a intervenção de operadores, acelerando desta forma a aquisição de informação. Pretende-se a restituição de um conjunto de objetos em simultâneo.

1.1 Enquadramento

O IGeoE faz parte da estrutura orgânica do Exército e tem por missão "...prover com informação geográfica o Exército, os outros ramos das Forças Armadas e a comunidade civil, assegurando a execução de atividades relacionadas com a ciência geográfica, a técnica cartográfica e a promoção e desenvolvimento de ações de investigação científica e tecnológica, no domínio do apoio geográfico e da geomática."¹.

¹in <http://www.igeoe.pt/index.php?id=28>[consultado em 31-01-2014]

Em termos de organização interna este é composto por quatro centros e uma repartição:

- Centro de Produção Cartográfica;
- Centro de Documentação Geográfica Militar;
- Centro de Desenvolvimento e Gestão de Informação;
- Centro de Formação Geográfica;
- Repartição de Apoio Geral.

O Centro de Produção Cartográfica tem a missão de "...produzir e manter a base de dados geográficos de todas as séries cartográficas à responsabilidade do IGeoE"², estando organizado em três departamentos:

- Departamento de Aquisição de Dados;
- Departamento de Processamento de Dados;
- Departamento de Saída de Dados.

O Departamento de Aquisição de Dados (DAD) tem a responsabilidade de "Apoiar, reconhecer e restituir as fotografias aéreas e/ou imagens digitais que servem de suporte à aquisição de informação para a base de dados geográficos, salvaguardando e preservando os processos correspondentes que originaram os dados adquiridos."³.

Neste departamento insere-se a Secção de Fotogrametria (SFot) que executa a restituição da informação presente nas fotografias digitais, em ambiente estereoscópico.

Periodicamente são realizados voos fotográficos sobre o Território Nacional, cujo planeamento e execução é da responsabilidade de diversos Organismos da Administração Pública. O principal intuito é a recolha de informação através de fotografia aérea, conforme o exemplo apresentado na figura 1.1.

Nestas circunstâncias, as aeronaves são equipadas com câmaras aéreas digitais, sendo as fotografias utilizadas para dois campos genéricos de aplicação:[Red07, p. 37]

- Fotointerpretação, que visa a análise e processamento de informação principalmente qualitativa;
- Fotogrametria, que visa a análise e processamento de informação geométrica e quantitativa.

Nesta dissertação a informação presente na fotografia aérea é utilizada no campo da fotogrametria, com objetivo de restituir a informação geométrica dos edifícios. Conforme já foi referido, pretende-se extrair o polígono relativo a cada edifício, e este deve ser adaptado à área do edifício de modo a representar a forma mais próxima da real. Na secção

²in <http://www.igeoe.pt/index.php?id=31#conteudo>[consultado em 20-01-2014]

³in <http://www.igeoe.pt/index.php?id=31#conteudo>[consultado em 20-01-2014]



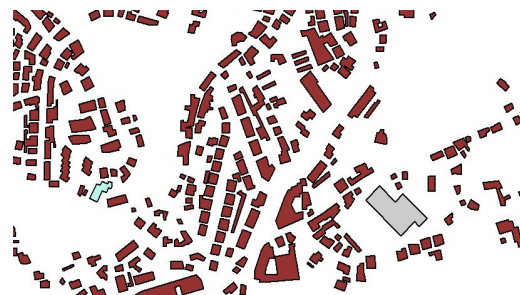
Figura 1.1: Fotografia aérea

2.3, irão ser abordados os procedimentos utilizados no IGeoE para efetuar a restituição de edifícios.

Observa-se na figura 1.2, a quantidade e variedade de polígonos existentes, numa área⁴ do Território Nacional.



(a) Numa Fotografia aérea



(b) Num SIG

Figura 1.2: Representação da informação dos polígonos

Na figura 1.2(b), podemos verificar que os polígonos resultantes, da restituição, apresentam uma grande variedade de formas e tamanhos.

Outro dos fatores que é importante referir, prende-se com as fotografias utilizadas nesta dissertação. Estas, não possuem nenhum tipo de processamento adicional, encontrando-se neste momento a ser utilizadas na cadeia de produção, no processo de restituição.

⁴Zona residencial da área metropolitana do Porto

1.2 Motivação

O processo até aqui descrito é atualmente utilizado pelo IGeoE e por grande parte das organizações que produzem cartografia, nacional e internacionalmente. O mesmo representa um elevado consumo de recursos humanos e materiais.

Em termos de recursos humanos é necessário considerar os custos de formação inicial, a experiência do operador e a remuneração de cada elemento.

Em termos de recursos materiais é necessário contabilizar, a aquisição de terminais com características e capacidade para executar as aplicações, a aquisição de licenças do sistema operativo e a aquisição de licenças do software de restituição, (para cada um dos terminais). Este último, acaba por ser o mais relevante, devido à especificidade deste tipo de trabalho e ao facto de poucas empresas desenvolverem aplicações nesta área elevando o custo de cada licença.

No processo de restituição, a aquisição de edifícios consome uma parte significativa dos recursos alocados ao processo. Assim, a automatização irá trazer uma evolução na forma de trabalhar potenciando os recursos existentes. Para estabelecer uma ordem de grandeza, atualmente a restituição de uma Carta Militar pode demorar entre seis a oito semanas de trabalho, em que sensivelmente 50% deste tempo é dedicado à aquisição de edifícios.

No mercado já existem aplicações que permitem solucionar este problema. No entanto estas apresentam algumas dificuldades no tratamento dentro do contexto do Território Nacional, devido à heterogeneidade dos edifícios. Outro dos inconvenientes prende-se com os dados de entrada, que estes utilizam, como por exemplo modelos digitais, pois não se encontram disponíveis na cadeia de produção.

1.3 Objetivos

O objetivo desta dissertação é o de proporcionar um método capaz de detetar e extrair edifícios, com informação altimétrica, para posterior inserção num Sistema de Informação Geográfica (SIG).

Pretende-se que o método seja utilizado após a fase de georreferenciação das operações preliminares da cadeia de produção cartográfica, conforme apresentado na figura 1.3, sendo introduzido entre a aerotriangulação e a restituição dos objetos, permitindo assim que os operadores, ao iniciarem o seu trabalho, disponham já de um conjunto de objetos levantados.

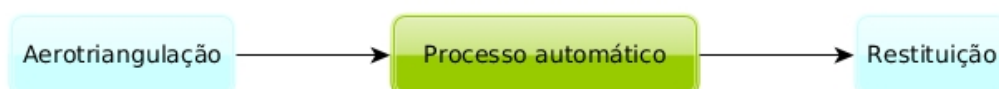


Figura 1.3: Localização do método a implementar

Desta forma o método terá, como dados de entrada, os produtos da aerotriangulação, sendo estes alvo de explicação no capítulo 2.

Como objetivo concreto, pretende-se extrair o contorno da parte superior dos edifícios, identificando um conjunto de características das imagens que formam os modelos estereoscópicos e que possam garantir a presença do edifício.

Para além da informação radiométrica presente em cada uma das fotografias, a aplicação dos fundamentos de fotogrametria analítica, com recurso às equações de colinearidade, permite determinar a informação altimétrica relevante para a definição dos contornos exteriores dos edifícios, ao nível da cobertura, sob a forma de um polígono.

1.4 Contribuições esperadas

Pretende-se acelerar o processo de restituição da cadeia de produção, proporcionando um método automático ou semi-automático de deteção de edifícios.

O método não inclui a utilização de modelos digitais de superfície, como dados de entrada, por não se encontrarem disponíveis nesta fase do processo. No entanto, poderá ser útil recorrer à informação local de elevações para a contextualização de áreas de interesse.

Com as coordenadas foto dos vértices que definem os polígonos, nas imagens do modelo estereoscópico, aplicar-se-ão as equações de colinearidade para o cálculo das coordenadas 3D de todos os vértices dos polígonos detetados.

A implementação deste método permitirá, eventualmente, criar uma base de conhecimento, para que esta seja aplicada à deteção de outros objetos incluídos no catálogo do IGeoE.

1.5 Organização do documento

Este documento encontra-se dividido em cinco capítulos sendo o primeiro dedicado à introdução do tema, a motivação de levou à sua elaboração, os objetivos e contribuições esperadas.

O segundo capítulo pretende dar uma visão mais detalhada sobre o processo de restituição utilizado no IGeoE, fazendo ainda a introdução ao tema da determinação altimétrica de pontos a partir do modelo estereoscópico.

O terceiro capítulo faz referência a um conjunto de trabalhos relacionados e que apresentam ter características importantes para a discussão, assim como um conjunto de técnicas relevantes que foram utilizadas na elaboração deste trabalho.

No quarto capítulo detalhamos a implementação do trabalho elaborado para solucionar o problema.

No quinto capítulo é feita uma avaliação à implementação, assim como à seleção dos valores dos parâmetros mais relevantes.

Para finalizar, o sexto capítulo contém as conclusões e considerações obtidas no decorrer do trabalho. Este incluído um conjunto de tarefas passíveis de estender este trabalho, e alternativas às utilizadas na implementação.



Restituição de objetos

Neste capítulo é abordado um conjunto de temáticas necessárias para a restituição de objetos. Conforme já foi referido brevemente no capítulo anterior, existe um conjunto de especificidades que necessitam de ser entendidas para melhor compreender o trabalho que se pretende realizar.

Inicialmente, fazemos uma breve abordagem sobre os dispositivos que permitem a aquisição fotográfica.

Em seguida são abordadas as fases das operações preliminares. Estas são necessárias para adquirir e preparar as fotografias, para que seja possível restituir objetos com precisão.

Segue-se a apresentação do processo de restituição, onde são apresentadas as normas que se encontram em vigor no IGeoE, os equipamentos e as considerações mais relevantes sobre o processo de restituição de objetos.

No final iremos abordar a utilização das equações de colinearidade, como forma de determinar a informação altimétrica. Apesar deste conjunto de equações não fazer parte do processo de restituição de objetos, as mesmas serão utilizadas na obtenção da altimetria, nesta dissertação.

2.1 Câmaras fotográficas

Para a obtenção das fotografias, as câmaras são montadas na fuselagem da aeronave e pode recorrer-se a várias lentes. O objetivo poderá ser a obtenção de imagens por bandas espectrais ou espaciais. Na figura¹ 2.1, podemos visualizar fotografias adquiridas no espectro do visível. Aplicando alterações do peso ou bandas das componentes da cor,

¹<http://www.oneonta.edu/faculty/baumanpr/geosat2/RS-Introduction/RS-Introduction.html>

na figura 2.1(a), é possível realçar determinadas características, conforme se encontra apresentado na figura 2.1(b), denominando-se esta técnica de falsa cor[frisp14].

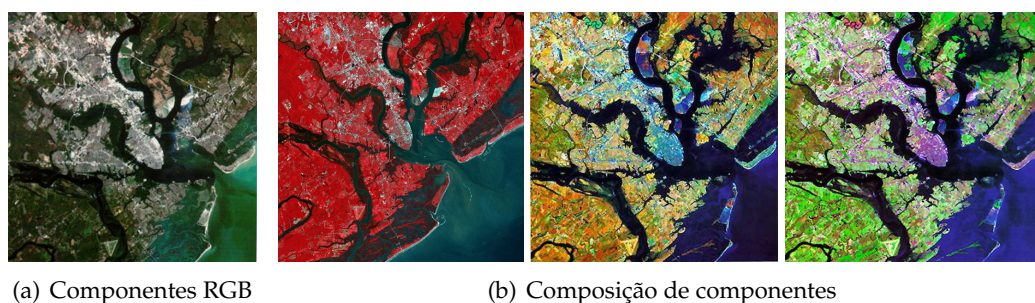


Figura 2.1: Fotografias do espectro visível

As fotografias utilizadas no IGeoE, atualmente foram adquiridas com a câmara **UltraCam Eagle** que está representada na figura^{2 3} 2.2. Esta tem a possibilidade de adquirir fotografias em vários espectros, contudo o IGeoE atualmente só tem disponíveis fotografias do espectro visível.



Figura 2.2: Câmara UltraCam Eagle

Anteriormente eram utilizadas câmaras analógicas, ficando a informação guardada em filme, sendo necessário efetuar um conjunto de procedimentos para revelar e digitalizar a fotografia. Com a introdução das câmaras digitais, no início do século, estes procedimentos deixaram de ser necessários, diminuindo o tempo entre a obtenção e utilização da informação e aumentando a qualidade. A utilização de fotografias proporciona uma afinidade entre o conteúdo da fotografia e os objetos que se convencionou representar na carta topográfica. As fotografias aéreas permitem ainda guardar uma elevada densidade de informação e tornam possível efetuar análises por épocas[Red07, p. 2].

Na tabela 2.1 podem ser consultadas as características do equipamento frequentemente utilizado nas fotografias fornecidas ao IGeoE.

Em fotografia aérea, a câmara pode ter dois tipos de orientação:

- Nadiral ou vertical - se esta estiver a apontar diretamente para o terreno na vertical;
- Oblíqua - se tiver outra direção.

²fonte: <http://ultracam.wordpress.com/2011/12/01/first-ultracam-eagle-digital-aerial-cameras-delivered/>

³fonte: <http://www.cnurs.com/chan-pin/shenyang/>

Tabela 2.1: Características das Câmaras

	UltraCam Eagle	DMC01
Orientação	nadiral	nadiral
Distância focal	79.8mm	120mm
Largura	13080px	7680px
Altura	20010px	13824px
Tamanho imagem	1047 Mb	425Mb
Tamanho do píxel	5.2 μ m	12 μ m

2.2 Operações preliminares

Iremos agora apresentar todo o processo necessário para que seja possível, a um operador, efetuar a restituição de objetos. A figura⁴ 2.3 ilustra de forma esquemática este processo.

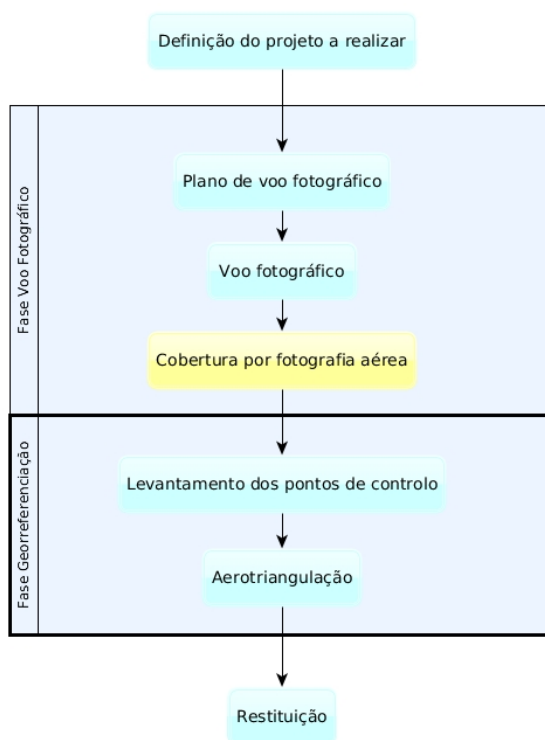


Figura 2.3: Diagrama de operações preliminares da cadeia de produção

O processo inicia-se com o levantamento de necessidades, tendo em consideração o grau de atualização das cartas ou inserido num plano superiormente, definido pela chefia do Exército.

Para efetuar a produção de cartografia, com o recurso a fotogrametria, é necessário dividir o processo em duas fases. Sendo a primeira o voo fotográfico e a segunda a

⁴adaptado: [Red10, pp. 3]

georeferenciação das fotografias tiradas durante o voo, segundo o referencial pretendido.

O planeamento da fase do voo fotográfico deverá ter em consideração um conjunto de aspetos, que iremos em seguida referir.

A escala da fotografia deve satisfazer os aspetos técnicos que permitam recolher a informação presente nas fotografias.

As sobreposições, medem-se em termos da percentagem de área comum representada em fotografias adjacentes. É necessário ter em atenção que o plano de voo terá que permitir a captura de um conjunto de fotografias existindo sobreposição longitudinal e lateral. Para a restituição de fotografias em ambiente estereoscópico a sobreposição longitudinal deve de ser de 60% e a lateral de 20 a 30% [Red07, p. 6], na figura⁵2.4 podemos ver uma representação deste tipo de sobreposição.

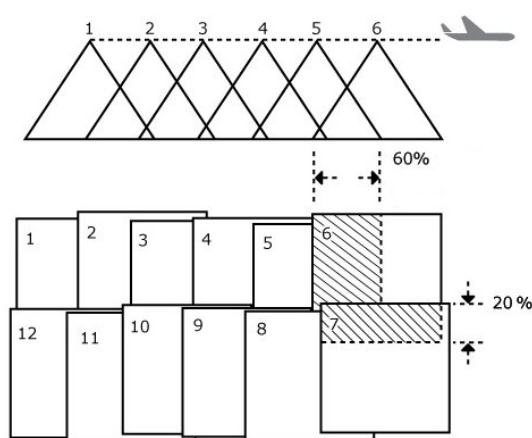


Figura 2.4: Sobreposição de fotografias

A objetiva utilizada ou disponível, influencia diretamente o grau de pormenor adquirido na fotografia, sendo necessário ajustar a altitude de voo para respeitar a escala da fotografia que se pretende obter.

O sensor das câmaras digitais influencia na maior e menor capacidade da recolha de informação. Sensores com maior número de píxeis por unidade de área permitem uma melhor definição da fotografia.

O traçado de voo é orientado para a cobertura pretendida. Se a área for extensa opta-se pelo modelo de faixas paralelas, conforme está presente na figura⁶ 2.5(a). Se a cobertura for uma faixa estreita opta-se pelo modelo da figura 2.5(b).

A época em que o voo se irá realizar terá de ter em atenção que a cobertura não deve ter obstáculos. As nuvens são um dos principais fatores a ter em consideração, pois a sua presença não permite a observação do terreno. O período do dia a que as fotografias são tiradas também pode afetar a qualidade, se os objetos projetarem muita sombra o processo de restituição é afetado.

⁵ fonte: <http://www.nrcan.gc.ca/earth-sciences/geomatics/satellite-imagery-air-photos/air-photos/about-aerial-photography/9687>

⁶ fonte: [Red07, p. 12]

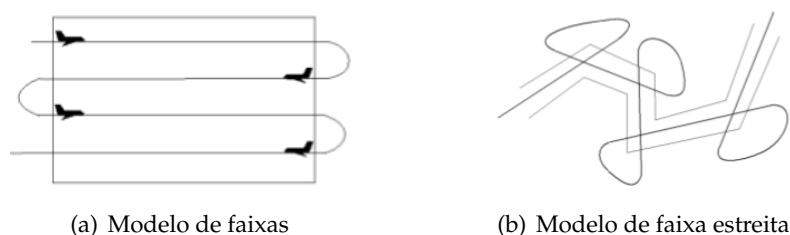


Figura 2.5: Traçado do voo

O tipo de filme/sensor também é levado em consideração, pois permitem capturar informação para diferentes objetivos. Os tipos de filme/imagem final mais comuns são:

- Infravermelho com cor;
- Infravermelho;
- Cor;
- Pancromático.

Outros aspetos a ter em consideração para o planeamento do voo são ainda o tipo de aeronave, a qual deve possuir características próprias para o efeito pretendido, como a estabilidade e a velocidade. O mapa de voo, onde está representado o esquema de navegação, e o próprio orçamento, também são estabelecidos nesta fase.

A georreferenciação e orientação, das imagens, constituem a segunda fase das operações preliminares da cadeia de produção cartográfica, conforme se encontra descrito na figura 2.3.

Este processo tem por objetivo a localização geográfica do centro de projeção de cada fotografia e a respetiva orientação espacial através da atribuição das suas coordenadas e ângulos de rotação segundo um determinado referencial. Neste caso as fotografias são referenciadas no sistema de coordenadas WGS84 militar.

Inicia-se o processo com o levantamento das coordenadas de pontos fotogramétricos e de controlo pelas equipas de topografia que se deslocam ao terreno. Estas selecionam um conjunto de pontos dispersos que sejam de inequívoca identificação nas fotografias. Como se pode verificar na figura 2.6, a informação do ponto contém um extrato da edição da carta anterior, um extrato da fotografia aérea, uma fotografia do local e as coordenadas do ponto. As coordenadas são obtidas através do sistema GNSS, com correção em tempo real de posicionamento da rede SERVIR.

Durante o processo de aerotriangulação⁷ os pontos de controlo são fundamentais para determinar a qualidade de georreferenciação. Uma boa georreferenciação visa a localização exata dos objetos e a sua correta relação.

⁷Para o processo de aerotriangulação, no IGeoE, é utilizada a aplicação *ImageStation Automatic Triangulation* (ISAT)



Figura 2.6: Informação de um ponto de controlo

A aerotriangulação procede a um conjunto de operações matemáticas e estatísticas que permitem orientar as fotografias entre si e relativamente ao terreno, assim como determinar as coordenadas objeto/terreno de qualquer ponto do modelo. Na figura 2.7 pode ver-se (a verde) a orientação de várias fotografias do bloco e (a azul) os pontos fotogramétricos utilizados.

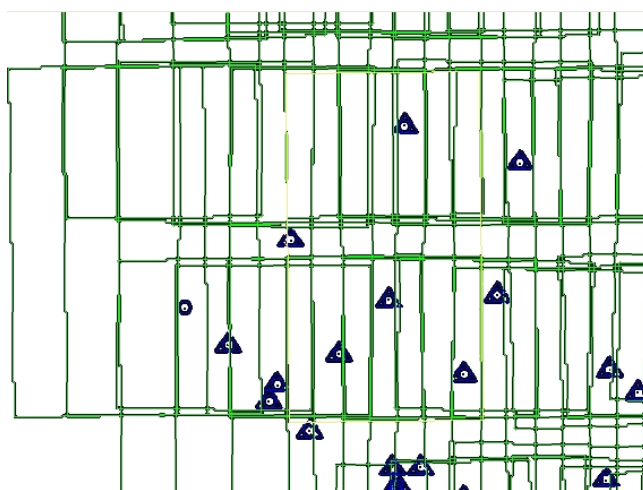


Figura 2.7: Pontos fotogramétricos e orientação das fotografias

O resultado da aerotriangulação pode ser ainda utilizado para produzir o modelo digital de superfície (MDS) e as ortofotos. No caso do IGeoE estes não são utilizados na SFot, pois toda a aquisição para o formato vetorial é efetuada a 3D.

Os MDS são criados quando existe necessidade. A sua realização implica a criação de uma nuvem de pontos, gerada na aplicação *ImageStation Automatic Elevation* (ISAE). Esta gera pontos 3D com um espaçamento previamente definido⁸.

As ortofotos são produzidas na aplicação **ImageStation OrthoPro** e servem de apoio

⁸Normalmente o espaçamento é de 10 metros, podendo ser reduzido

aos restantes processos na produção da Carta Militar.

Não se antevê viabilidade em utilizar estes produtos, no contexto desta dissertação, devido ao MDS não ter um espaçamento suficientemente pequeno para permitir a sua utilização na deteção de edifícios. Assim como a utilização de ortofotos, por corresponderem a uma projeção ortogonal a 2D da informação das fotografias, e consequentemente não permitir a determinação da cota dos edifícios. Na secção 2.2.1, o processo de aerotriangulação será detalhado com maior pormenor.

Neste momento é possível iniciar-se a restituição dos objetos para a produção de cartografia. Esta consiste em extrair do modelo estereoscópico a informação de interesse para a produção da Carta Militar. Esta encontra-se definida no guia de extração e completagem que contempla os seguintes temas:

- Rede viária;
- Altimetria;
- Edifícios e Construções;
- Hidrografia;
- Vegetação.

Na figura 2.8(a) podemos observar um operador, na sua estação de trabalho, sendo representada uma secção restituída na figura 2.8(b). Este conjunto de procedimentos serão abordados na secção 2.3, com maior detalhe.



(a) Operador a restituir



(b) Restituição

Figura 2.8: Procedimentos de restituição

2.2.1 Aerotriangulação

Conforme já foi referido, para efetuar a aerotriangulação no IGeoE, utiliza-se a aplicação ISAT. Esta tem como objetivo o refinamento dos parâmetros de orientação externa das fotografias aéreas para posterior utilização na restituição dos vários objetos.

Na figura 2.9 é apresentado um diagrama com os dados de entrada e saída de cada uma das aplicações utilizadas. Para efetuar esta tarefa são necessárias as fotografias e um

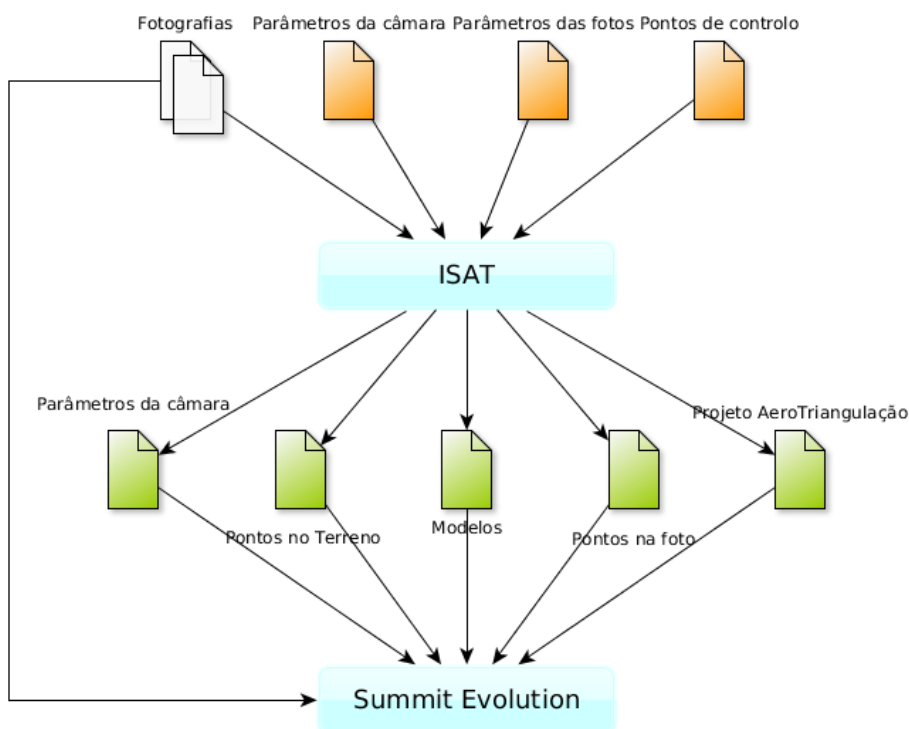


Figura 2.9: Diagrama de entrada e saída de dados das aplicações ISAT e Summit Evolution

conjunto de dados de entrada. Estes encontram-se identificados na figura 2.9, a amarelo.

O ficheiro dos **parâmetros da câmara** contem a informação da orientação interna, a qual permite construir o feixe de raios utilizado no momento da exposição[Red07, p. 65], e inclui:

- Distância focal calibrada da câmara;
- Posição do ponto principal (PP) no negativo ou sensor;
- Os parâmetros de distorção do sistema de lentes fotográfico.

O ficheiro dos **parâmetros da fotografia** contém a informação da orientação externa, relativa à posição da máquina no momento de exposição, incluindo:

- Coordenadas do centro de projeção;
- Atitude da câmara, em valores angulares.

Conforme podemos verificar na figura 2.10 os parâmetros de orientação interna encontram-se no espaço imagem, enquanto os parâmetros de orientação externa localizam-se no espaço objeto.

O ficheiro dos **pontos de controlo** contém as coordenadas dos pontos de controlo que as equipas de topografia levantaram no terreno com a utilização da tecnologia GNSS.

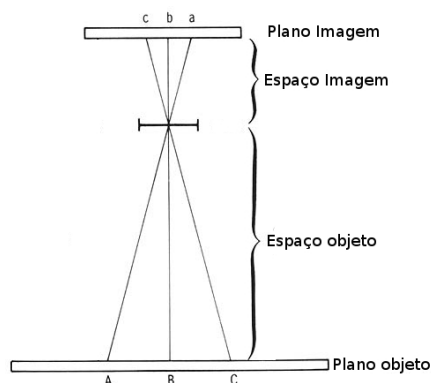


Figura 2.10: Localização das orientações interna e externa

Do conjunto total de pontos de controlo são escolhidos pontos para serem referenciados manualmente em todas as fotografias em que estejam presentes, (exemplo na figura 2.11). Estes designam-se por pontos fotogramétricos. Os restantes são utilizados, no final do processo para aferir a qualidade posicional do processo.



Figura 2.11: Identificação de pontos fotogramétricos em várias fotografias

Para além dos pontos inseridos manualmente, a aplicação identifica um conjunto de pontos homólogos, em ambas as fotografias. Os pontos são selecionados em zonas específicas, denominadas zonas de Von Gruber, garantindo que os pontos se encontram dispersos pela totalidade das fotografias. Na figura⁹ 2.12 encontra-se representada a mesma área, que é identificada em 6 fotografias diferentes.

Após a correlação dos pontos artificiais e inserção de pontos fotogramétricos, temos a possibilidade de determinar a orientação que uma fotografia tem em relação à outra. Este procedimento denomina-se orientação relativa, onde são determinadas as posições

⁹fonte: [Red10, p. 98]

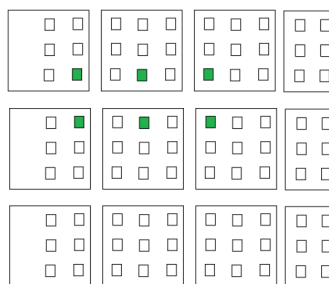


Figura 2.12: Zonas de Von Gruber a verde correspondem à mesma zona no objeto

e atitude angular relativa das fotografias do modelo. Assim, podem ser recriadas as condições do momento em que as fotografias foram tiradas, fazendo com que os raios projetivos dos pontos artificiais e fotogramétricos se intersectem no espaço modelo [Gon06, p. 38].

Depois do modelo se encontrar orientado de forma relativa, a orientação absoluta tem por objetivo efetuar a transformação do sistema de coordenadas do modelo para coordenadas objeto[Gon06, p. 50], sendo assim possível ler coordenadas 3D.

Após a conclusão destas operações, o ISAT exporta o conjunto de ficheiros que se encontra na figura 2.9 a verde.

- Ficheiro de parâmetros da câmara - contém as orientações internas;
- Ficheiro de pontos no terreno - contém as coordenadas objeto dos pontos artificiais e fotogramétricos, assim como o erro associado a cada um deles;
- Ficheiro de pontos na foto - contém as coordenadas foto dos pontos artificiais e fotogramétricos, com a tolerância inerente a cada um deles e as orientações externas de cada fotografia;
- Ficheiro de modelos - contém a informação dos modelos com a orientações relativas e absolutas;
- Ficheiro de projeto de aerotriangulação - contém a informação de unidades e a tolerâncias das orientações;

2.3 Restituição

O processo de restituição é efetuado com a aplicação Summit Evolution da DAT/EM, que utiliza como dados de entrada as fotografias e os ficheiros gerados pelo ISAT.

O operador vai ter à sua disposição o ambiente estereoscópico, correspondente ao modelo que está a visualizar, e um conjunto de ferramentas de "desenho" para efetuar o seu trabalho. Com estas ferramentas pode desenhar vários elementos e alterar os seus atributos. Todos estes elementos são inseridos diretamente no SIG.

2.3.1 Normas

Para garantir uma homogeneidade na forma como a restituição é efetuada pelos vários operadores, o IGeoE possui um conjunto de Normas de Aquisição, tendo os operadores formação segundo estas normas.

Após a restituição dos objetos da carta em questão, a informação vetorial é entregue para a validação interna da secção de fotogrametria, onde se vai verificar o cumprimento das normas.

Segundo as normas técnicas[IGe08] para a aquisição de dados da série M888 do IGeoE, podem ser restituídos dois tipos de edifícios: casas regulares ou irregulares.

Considera-se como casa regular a "Construção destinada a habitação e afins, de formato poligonal e de lados ortogonais"[IGe08] e são restituídas pelos vértices do edifício.

A casa irregular é uma "Construção de carácter permanente, destinada ou não à habitação e que se caracteriza pela existência de uma cobertura suportada por paredes"[IGe08]. Estas são restituídas pelos vértices. Se tiverem arcos de circunferência, devem acompanhar o contorno do edifício pelo beiral.

A escala 1:25 000 é a maior escala para a qual o IGeoE produz cartografia. Desta forma, nem todos os edifícios presentes nas fotografias são restituídos, por não representarem informação relevante. Adquirem-se individualmente os edifícios com as seguinte características:

- Separação horizontal superior a 3m;
- Construções para estacionamento com altura superior a 3m;
- Construções que tenham campos de jogos ou piscinas na parte superior.

Não são adquiridos:

- Pilares e cobertura se não houver paredes laterais;
- Pequenas construções para guardar ferramentas agrícolas ou motores de rega;
- Construções com menos de 3 m×3 m de área ou equivalente;
- Construções que estejam sobre construções;
- Construções que façam parte de uma muralha.

Se existir a necessidade de restituir edifícios contíguos, é necessário que estes sejam agrupados no mesmo polígono. Em zonas urbanas e históricas este procedimento é mais frequente, devido à densidade de edifícios.

2.3.2 Equipamentos

Para efetuar o processo de restituição, o operador trabalha num terminal com capacidade de processar a informação dos modelos estereoscópicos. Isso implica que tenham uma

elevada capacidade de processamento ao nível da placa gráfica (*Graphics Processing Unit* (GPU)).

Este terminal funciona em combinação com um equipamento ocular composto por duas lentes que, em sincronização com a placa gráfica, bloqueia uma lente de cada vez, mostrando no ecrã a fotografia correspondente ao olho não bloqueado. Na figura 2.13(a) podemos ver um deste equipamentos.



Figura 2.13: Equipamentos utilizados na restituição

Para manipular a aplicação é utilizado um rato 3D, apresentado na figura 2.13(b). Este não necessita de ser movimentado na superfície da mesa e permite:

- Navegar nas opções da aplicação;
- Movimentar uma das fotografias, para eliminar a paralaxe;
- Movimentar o cursor em X e Y;
- Alterar as escalas de visualização.

2.3.3 Processo

O processo de restituição divide-se em duas partes. Na primeira o operador tem que percorrer o modelo estereoscópico identificando os objetos visualmente. Na segunda este procede à vetorização do objeto, respeitando as normas aprovadas superiormente.

Mesmo sendo relativamente fácil identificar os objetos com o olho humano, por vezes surgem dúvidas. Para o seu esclarecimento, o operador pode recorrer à carta topográfica da edição anterior, ou colocar uma dúvida, a ser esclarecida durante os trabalhos de campo.

Para a vetorização, o operador delimita cada um dos objetos, no caso de objetos do tipo área. No caso de objetos lineares são colocados vértices de modo a definir a geometria ao longo de uma linha. Nos objetos representados por um ponto, é colocado um ponto com os atributos adequados no centro do objeto. O operador tem a preocupação constante de manter a paralaxe eliminada, enquanto coloca o cursor sobre o objeto permitindo assim adquirir as coordenadas 3D de todos os vértices para a base de dados.

Na figura 2.14(a) podemos ver uma parte da fotografia onde ainda não foi restituído nenhum objeto, estando a fotografia sem nenhuma informação vetorial. Na figura 2.14(b)

o operador já identificou e assinalou a informação, na figura 2.14(c) está representado somente o trabalho realizado. A figura 2.14(d) apresenta a visualização a 3D da zona de terreno.

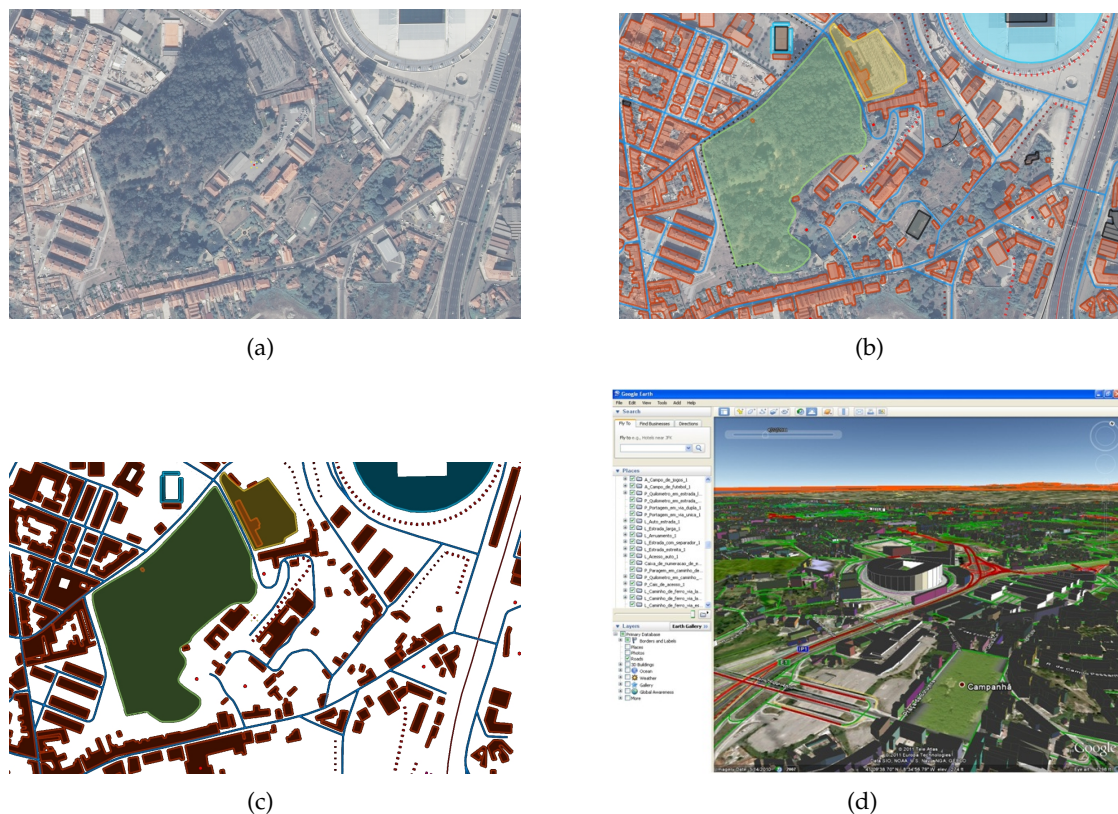


Figura 2.14: Processo de restituição

Como foi referido na secção 2.3.1, para os edifícios, quanto mais complexa for a arquitetura da parte superior do edifício, maior será o número de pontos a adquirir.

Se consideramos um edifício cuja forma seja a de um paralelepípedo, a sua secção superior é um retângulo. Neste caso o operador tem que marcar os quatro vértices, para restituir o edifício. Se levarmos em consideração que na carta número 122 do Porto, foram necessários restituir cerca de 53 000 edifícios, o número de vértices que o operador tem que identificar é relativamente elevado.

A restituição de todos os objetos de uma carta pode demorar seis a oito semanas. No caso de uma atualização normalmente são despendidas seis semanas. Se houver necessidade de refazer toda a carta, o processo poderá demorar oito semanas.

A decisão de refazer as cartas depende da qualidade dos dados e equipamentos utilizados, na altura da sua elaboração, em comparação com os disponíveis no presente. Este facto é relevante, pois de momento existe informação que foi restituída a partir de fotografias digitais, de fotografias analógicas que foram digitalizadas e de fotografias analógicas¹⁰. Se existir uma diferença substancial na qualidade da informação anterior,

¹⁰Estas permitiam uma escala média de 1:12 500 através de zoom ótico

para pior opta-se por refazer a carta.

O processo de restituição de edifícios, representa em média, 50% do trabalho necessário para produzir uma carta. Em cartas que estejam localizadas junto ao litoral este valor aumenta substancialmente, assim como também diminui quando as cartas se localizam no interior do país.

Em termos de erro inerente à aquisição de edifícios é necessário ter em atenção o erro de graficismo. Se consideramos que a finalidade do processo é a produção da carta militar na escala 1:25 000, podemos utilizar o limite de perceção visual em papel, para definir o erro de graficismo admissível na carta. Este limite refere que não é possível distinguir um segmento de um ponto, se o segmento tiver um comprimento igual ou inferior a 0.2 mm[Pes06, cap2 p.5]. Nesta escala o valor de erro seria 5 metros no terreno.

O IGeoE como entidade produtora de informação geográfica, encontra-se obrigada a respeitar os compromissos internacionais, nomeadamente os documentos ratificados no âmbito da *North Atlantic Treaty Organization* (NATO). O documento *Standardization Agreement* (STANAG)[22102] aplica uma classificação aos dados geográficos, em termos de precisão horizontal e vertical, mediante a escala em que os dados são utilizados. A classificação mais elevada obriga, no caso da escala 1:25 000, a uma precisão horizontal de 12.5 m. Enquanto a precisão vertical deve ser inferior a 2.5 m. Sendo este valor mais permissivos que os apresentados anteriormente.

Se recorremos às normas da DGT, a autoridade Nacional em termos de cartografia, estas determinam que, para a escala de 1:10 000, para se obter conformidade posicional o erro quadrático médio (EMQ) deve ser inferior a 1.5 m, sendo que numa amostra representativa, 90% dos pontos deve apresentar valores inferiores a 2.3 m [dT13, p. 24]. Em altimetria o EMQ é de 1.8 m com 90% dos pontos com valores inferiores a 3 m [dT13, p. 19].

No entanto, é necessário ter em atenção que a qualidade posicional da informação produzida no IGeoE, possui erros inferiores aos referidos anteriormente, relativos à escala 1:25 000. No controlo de qualidade efetuado nos blocos mais recentes o erro médio foi inferior a 2 m no terreno, em planimetria e altimetria. Desta forma, os objetos restituídos, para a base de dados geográficos, possuem uma qualidade posicional de uma escala 1/10 000.

Na tabela 2.2 são apresentados os valores de erro, mediante a escala utilizada. Para efetuar o cálculo do número de píxeis, foi utilizado o *ground sample distance* (GSD), das fotografias do bloco que pretendemos utilizar. Estas possuem um GSD de 0.52 m.

Tabela 2.2: Relação entre escalas e erros em planimetria

Escala	erro de graficismo	erro no terreno	erro em píxeis
1:25 000	0.2 mm	5 m	≈ 10
1:25 000 NATO	0.5 mm	12.5 m	≈ 25
1:10 000	0.2 mm	2 m	≈ 4
1:10 000 DGT	0.15 mm	1.5 m	≈ 4

2.4 Equações de colinearidade

Conforme foi referido, pretende-se recorrer às equações de colinearidade fazendo com estas uma intersecção direta espacial para determinar a informação altimétrica.

Nos dados fornecidos pelo programa ISAT, está disponível a informação da orientação externa, localizada no ficheiro pontos na foto, apresentado na figura 2.9. Estes valores são determinados através de um conjunto de sensores que a câmara dispõe.

Em fotogrametria, o movimento contempla o seu próprio formalismo, em termos de rotação em torno de um dos eixos do sistema, designadamente:

- omega Ω - rotação em torno do eixo x;
- phi ϕ - rotação em torno do eixo y;
- kappa κ - rotação em torno do eixo z.

Esta é a nomenclatura utilizada no manual [Red10, p. 54], mas é usual encontrar bibliografia com outras designações [EWb].

Com os valores da atitude da câmara irá proceder-se à rotação do plano da fotografia para um plano paralelo ao plano xy, através da matriz de rotação M apresentada na equação 2.1.

$$M = \begin{bmatrix} \cos \Omega \cos \kappa & -\cos \phi \sin \kappa & \sin \phi \\ \cos \Omega \sin \kappa + \sin \Omega \sin \phi \cos \kappa & \cos \Omega \cos \kappa - \sin \Omega \sin \phi \sin \kappa & -\sin \Omega \cos \phi \\ \sin \Omega \sin \kappa - \cos \Omega \sin \phi \cos \kappa & \sin \Omega \cos \kappa + \cos \Omega \sin \phi \sin \kappa & \cos \Omega \cos \phi \end{bmatrix} \quad (2.1)$$

Para simplificar a identificação dos elementos da matriz, na formulação das restantes equações, deste capítulo, iremos aplicar a seguinte nomenclatura:

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \quad (2.2)$$

A figura¹¹ 2.15 representa a existência de colinearidade, no momento em que fotografia foi tirada, entre o centro de projecção, ponto O, o ponto objeto P, e o ponto foto p, devido ao principio de que a luz se propaga em linha reta.

¹¹ fonte:[Red10, p. 58]

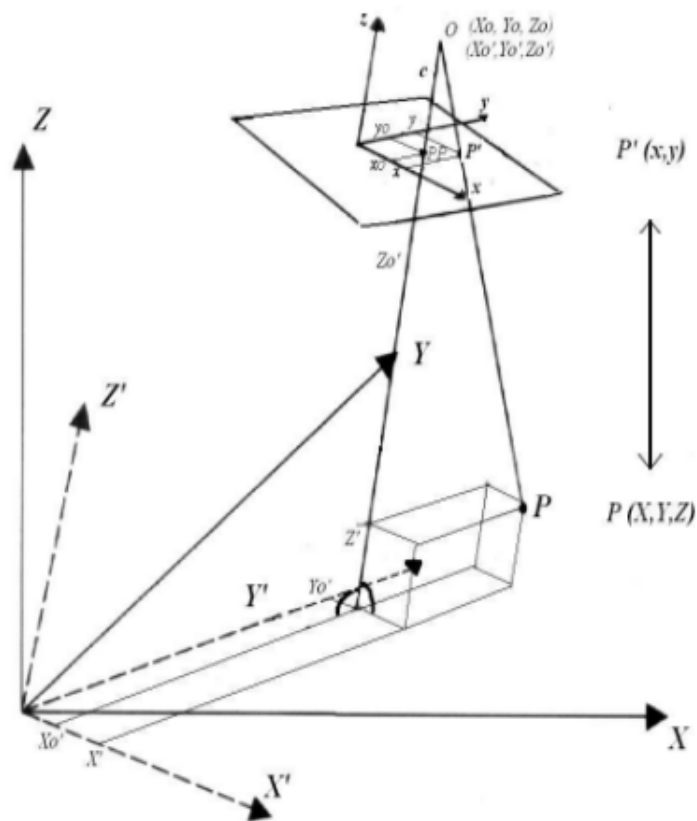


Figura 2.15: Sistema de eixos onde está representada a colinearidade

Em termos de legenda, a figura 2.15 apresenta um conjunto de valores que neste documento têm o seguinte significado:

- x, y, z - sistema de coordenadas 3D foto com origem no ponto médio da fotografia;
- X, Y, Z - sistema de coordenadas objeto;
- $x_o, y_o, 0$ - coordenadas do ponto principal da fotografia;
- x_o, y_o, c - coordenadas do centro de projeção;
- c - Distância focal.

As equações de colinearidade permitem a relação entre coordenadas foto e coordenadas objeto, e são expressas conforme se encontra no sistema de equações 2.3:

$$\begin{cases} x = x_o - c \frac{m_{11}(X - X_o) + m_{21}(Y - Y_o) + m_{31}(Z - Z_o)}{m_{13}(X - X_o) + m_{23}(Y - Y_o) + m_{33}(Z - Z_o)} \\ y = y_o - c \frac{m_{12}(X - X_o) + m_{22}(Y - Y_o) + m_{32}(Z - Z_o)}{m_{13}(X - X_o) + m_{23}(Y - Y_o) + m_{33}(Z - Z_o)} \end{cases} \quad (2.3)$$

Estas podem ser expressas em ordem às coordenadas objeto, conforme se encontra no sistema de equações 2.4:

$$\begin{cases} X = X_o + (Z - Z_o) \frac{m_{11}(x-x_o) + m_{12}(y-y_o) - m_{13}.c}{m_{31}(x-x_o) + m_{32}(y-y_o) - m_{33}.c} \\ Y = Y_o + (Z - Z_o) \frac{m_{21}(x-x_o) + m_{22}(y-y_o) - m_{23}.c}{m_{31}(x-x_o) + m_{32}(y-y_o) - m_{33}.c} \end{cases} \quad (2.4)$$

Assim com o sistema de equações 2.4, é possível determinar, para um ponto da fotografia (x, y) , o valor X, Y objeto, se tivermos o valor de Z previamente disponível. O valor de Z pode ser obtido a partir de um MDS ou através de um processo de cálculo. Devido ao facto de não existir um MDS disponível entre os processos de aerotriangulação e restituição, determinamos os valor do Z com o auxilio da intersecção direta, com uma aproximação inicial, seguindo-se a aplicação do método dos mínimos quadrados (MMQ) para otimizar o resultado.

A figura¹² 2.16 apresenta graficamente o processo de intersecção direta. Este consiste em determinar as coordenadas objeto de um ponto a partir das respetivas coordenadas foto, em ambas as fotografias (x_1, y_1) e (x_2, y_2) do modelo estereoscópico, referentes ao mesmo ponto objeto, conhecendo os parâmetros de orientação interna e externa das duas fotografias. Deduzindo as equações de colinearidade [Red10, p. 70] expressas em ordem às coordenadas objeto, obtemos as seguinte equações 2.5.

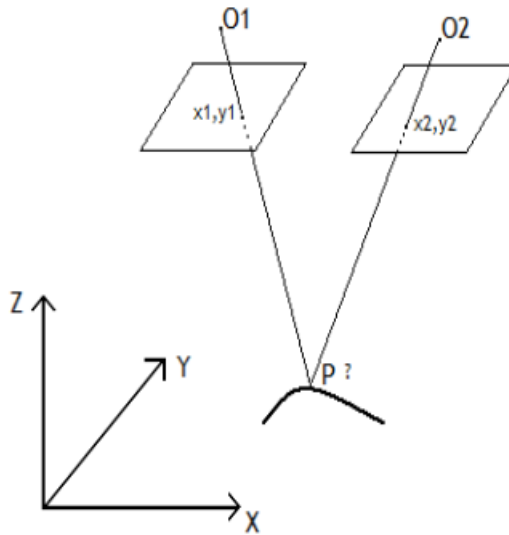


Figura 2.16: Intersecção direta espacial

¹²fonte:[Red10, p. 70]

$$\left\{ \begin{array}{l} X = X_{o1} + (Z - Z_{o1})K_{x1} \\ Y = Y_{o1} + (Z - Z_{o1})K_{y1} \\ X = X_{o2} + (Z - Z_{o2})K_{x2} \\ Y = Y_{o2} + (Z - Z_{o2})K_{y2} \end{array} \right. \quad \text{onde} \quad \left\{ \begin{array}{l} K_x = \frac{m11(x-x_o)+m12(y-y_o)-m13.c}{m31(x-x_o)+m32(y-y_o)-m33.c} \\ K_y = \frac{m21(x-x_o)+m22(y-y_o)-m23.c}{m31(x-x_o)+m32(y-y_o)-m33.c} \end{array} \right. \quad (2.5)$$

O sistema de equações 2.5 permitem determinar as coordenadas aproximadas de um ponto objeto, ficando as quatro equações lineares com três incógnitas X,Y e Z[Red10, p. 71]. Utilizando a primeira e terceira equação podemos determinar Z, pela equação 2.6.

$$Z = \frac{X_{o2} - Z_{o2} \cdot K_{x2} + Z_{o1} \cdot K_{x1} - X_{o1}}{K_{x1} - K_{x2}} \quad (2.6)$$

São utilizadas a primeira e terceira equações, sendo também possível utilizar a segunda e quarta equações.

É necessário ter em atenção que esta é a primeira aproximação a ter em conta para a determinação das coordenadas objeto. Após este passo são calculados os desvios necessários para a compensação, através do MMQ.

Através da linearização das equações de colinearidade, é possível determinar os parâmetros que não são conhecidos. Recorre-se neste caso a observações redundantes, para formular um sistema de equações de observação, e estimar os parâmetros pelo MMQ.

Para cada ponto na fotografia são determinadas duas equações de observação, neste caso particular utilizamos um modelo estereoscópico, com dois pontos conhecidos que irá dar origem a quatro equações e três incógnitas, dX, dY, dZ, conforme representado na equação 2.7.

$$\left\{ \begin{array}{l} vx_1 = \left(\frac{\partial x}{\partial X}\right)_1^{(0)} dX + \left(\frac{\partial x}{\partial Y}\right)_1^{(0)} dY + \left(\frac{\partial x}{\partial Z}\right)_1^{(0)} dZ + (x_1^{(0)} - \bar{x}_1) \\ vy_1 = \left(\frac{\partial y}{\partial X}\right)_1^{(0)} dX + \left(\frac{\partial y}{\partial Y}\right)_1^{(0)} dY + \left(\frac{\partial y}{\partial Z}\right)_1^{(0)} dZ + (y_1^{(0)} - \bar{y}_1) \\ vx_2 = \left(\frac{\partial x}{\partial X}\right)_2^{(0)} dX + \left(\frac{\partial x}{\partial Y}\right)_2^{(0)} dY + \left(\frac{\partial x}{\partial Z}\right)_2^{(0)} dZ + (x_2^{(0)} - \bar{x}_2) \\ vy_2 = \left(\frac{\partial y}{\partial X}\right)_2^{(0)} dX + \left(\frac{\partial y}{\partial Y}\right)_2^{(0)} dY + \left(\frac{\partial y}{\partial Z}\right)_2^{(0)} dZ + (y_2^{(0)} - \bar{y}_2) \end{array} \right. \quad (2.7)$$

Sendo $\left(\frac{\partial \dots}{\partial \dots}\right)_i^{(0)}$ as derivadas parciais em ordem a cada uma das variáveis, $x_i^{(0)}$ e $y_i^{(0)}$ as coordenadas calculadas pelas equações de colinearidade com base nos valores aproximados calculados em 2.5 e 2.6, para o primeiro ponto, e \bar{x}_i e \bar{y}_i os valores medidos na fotografia.

Este sistema pode ser resolvido através de cálculo matricial, recorrendo à equação apresentada na equação 2.8. A equação 2.9 representa a construção da matriz A com as derivadas parciais e na 2.11 encontra-se a construção da matriz l com as diferenças entre as coordenadas foto medidas e calculadas.

$$X = (A^T \cdot A)^{-1} \cdot A^T \cdot l \quad (2.8)$$

$$A = \begin{bmatrix} \frac{\partial x}{\partial X} 1 & \frac{\partial x}{\partial Y} 1 & \frac{\partial x}{\partial Z} 1 \\ \frac{\partial y}{\partial X} 1 & \frac{\partial y}{\partial Y} 1 & \frac{\partial y}{\partial Z} 1 \\ \frac{\partial x}{\partial X} 2 & \frac{\partial x}{\partial Y} 2 & \frac{\partial x}{\partial Z} 2 \\ \frac{\partial y}{\partial X} 2 & \frac{\partial y}{\partial Y} 2 & \frac{\partial y}{\partial Z} 2 \end{bmatrix} \quad (2.9)$$

As derivadas parciais já se encontram devidamente discriminadas em [Red10, p. 71] e são apresentadas na equação 2.10. Em que D representa o denominador das equações de colinearidade, N_x o numerador da primeira equação e N_y o numerador da segunda equação no sistema de equações 2.3.

$$\begin{aligned} \frac{\partial x}{\partial X} &= -\frac{c}{D^2}(D.m11 - N_x.m13) \\ \frac{\partial x}{\partial Y} &= -\frac{c}{D^2}(D.m21 - N_x.m23) \\ \frac{\partial x}{\partial Z} &= -\frac{c}{D^2}(D.m31 - N_x.m33) \\ \frac{\partial y}{\partial X} &= -\frac{c}{D^2}(D.m12 - N_y.m13) \\ \frac{\partial y}{\partial Y} &= -\frac{c}{D^2}(D.m22 - N_y.m23) \\ \frac{\partial y}{\partial Z} &= -\frac{c}{D^2}(D.m32 - N_y.m33) \end{aligned} \quad (2.10)$$

$$l = \begin{bmatrix} \overline{x_1} - x_1^{(0)} \\ \overline{y_1} - y_1^{(0)} \\ \overline{x_2} - x_2^{(0)} \\ \overline{y_2} - y_2^{(0)} \end{bmatrix} \quad (2.11)$$

Após o cálculo da equação 2.8, ficamos com um vector com desvios necessários para compensar as coordenadas. Este processo é iterativo até chegar a um valor de compensação pré-definido pela tolerância.

Para determinar a diferença entre o valor obtido pela aproximação inicial e o MMQ, foram selecionados 3 pontos do apoio de topografia, do bloco do Minho, para comparar as coordenadas entre ambos. Na tabela 2.3 podemos verificar a diferença entre as coordenadas calculadas com recurso às equações de colinearidade e as coordenadas aplicando o MMQ. Analisando esta tabela verificamos que o ajustamento do MMQ é mais significativo na componente do Z, efetuando ajustamentos na ordem dos centímetros.

Tabela 2.3: Determinação de coordenadas

ponto		Eq Colinearidade	Eq Colinearidade + MMQ	Diferença
199	X	183047.9344	183047.9355	-0.0011
	Y	509393.1862	509393.1876	-0.0014
	Z	425.9712	425.9562	0.0152
171	X	183463.3171	183463.3174	-0.0003
	Y	512278.3332	512278.3347	-0.0015
	Z	458.7198	458.7085	0.0113
203	X	182683.5774	182683.5775	-0.0001
	Y	506483.8761	506483.8757	0.0004
	Z	238.8162	238.8145	0.0017

Na tabela 2.4 podemos verificar a diferença entre o cálculo de coordenadas com as equações de colinearidade e MMQ e as coordenadas lidas por um operador na aplicação de restituição. Podemos ainda verificar que nas componentes de X, Y e Z os valores encontram-se abaixo do metro. De referir ainda que o facto do ponto 171 apresentar valores superiores aos restantes deve-se, a este ponto ter sido retirado num tanque de água, sendo nestas circunstâncias, em ambiente estereoscópico, relativamente difícil determinar a lateral do tanque.

Tabela 2.4: Diferença entre coordenadas calculadas e retiradas na aplicação Summit

ponto		Eq Colinearidade + MMQ	Summit	Diferença
199	X	183047.9355	183048.2030	-0.26
	Y	509393.1876	509393.3240	-0.11
	Z	425.9562	426.4000	-0.44
171	X	183463.3174	183463.7950	-0.48
	Y	512278.3347	512277.4400	0.89
	Z	458.7085	460.2750	-1.57
203	X	182683.5775	182684.0750	-0.51
	Y	506483.8757	506483.9400	-0.06
	Z	238.8145	238.2600	0.55



Trabalho Relacionado

Neste capítulo será apresentado um conjunto de trabalhos que, direta ou indiretamente, abordam alguns dos problemas apresentados por esta dissertação.

Como já foi referido, os dados de entrada para a elaboração da dissertação, já se encontram devidamente determinados. Alguns dos trabalhos aqui em ênfase utilizam dados de entrada que não temos disponíveis. Apesar de não ser possível a sua utilização, existe um conjunto de técnicas apresentadas que são importantes mencionar.

A forma de recolha de informação tem sido alterada ao longo do tempo, enquanto nas décadas de 80 e 90 o enfoque estava na deteção de edifícios em fotografia aérea, atualmente este passou para a utilização de imagem de satélite. Esta alteração deve-se ao elevado investimento e desenvolvimento tecnológico nos sensores presentes em órbita, sensores estes que permitem a aquisição periódica, com cobertura de grandes áreas a custos competitivos. No entanto este tipo de imagens apresenta algumas desvantagens no que diz respeito à dificuldade de aquisição de informação de forma precisa[Bal13].

Os métodos de deteção de edifícios, em fotografia aérea, funcionam fundamentalmente de dois modos, através de deteção de características(*footprint detection*) ou utilizando extração com base em probabilidades [Shi09].

A extração com base deteção de características necessitam de estabelecer um conjunto de condições, para representar a realidade. Se as condições não forem corretas, esta extração vai apresentar erros. Um destes casos é o facto de considerarmos que todos os edifícios possuem arestas visíveis. Desta forma não é possível detetar edifícios com linhas curvas e além disso construções como piscinas ou estradas serão (erradamente) consideradas edifícios.

No contexto deste capítulo quando nos referimos a telhado é necessário ter em atenção que nos referimos à parte superior de um edifício, podendo ser um telhado composto

de vários planos ou uma cobertura.

A execução baseada em probabilidades necessitam de um grande conjunto de dados de treino, para efetuarem uma classificação adequada. Este tipo de extração acaba por funcionar bem no caso em que os edifícios apresentam bastante semelhança. Mesmo assim os autores necessitam, de um elevado conjunto de treino para que seja possível classificar os restantes. Devido à grande diversidade das construções presentes em Portugal, este método não é o mais apropriado, pois era necessário despende muito recursos humanos para criar o conjunto de treino.

Nesta capítulo é abordada a forma como os diferentes espaços de cores, organizam a informação. Este detalhe é importante para a execução desta dissertação, pois diferentes espaços de cor, possuem métricas diferenciadas para o cálculo da distâncias entre elas.

São ainda abordadas duas técnicas de deteção de pontos característicos, em imagens. Estas iram permitir encontrar os pontos notáveis em cada uma das fotografias do modelo. Sendo em seguida utilizada uma técnica de correlação de pontos, para nos indicar aqueles que possuem correspondência.

3.1 Extração com base em características

Os métodos podem variar com diversos critérios, tais como a resolução, a complexidade das fotografias, o tipo de representação ou o tipo de sensores [May99].

Este conjunto de métodos podem efetuar a deteção diretamente na fotografia ou no processamento desta, pois a aplicação de algoritmos de processamento de imagem pode permitir evidenciar uma determinada característica que estamos a tentar encontrar.

Estas características podem ser vértices, arestas, linhas, sombras, paredes verticais, etc, conforme é possível consultar na tabela 3.1 para o caso dos artigos analisados.

Tabela 3.1: Características utilizadas

Artigo	Vértices	Arestas	Linhas	Sombras
[HN88]	x	x	x	x
[LN98]	x	x	x	x
[NLH97]	x	x	x	x
[ST10]	-	-	-	-
[GHS01]	-	x	x	-
[SU08]	x	x	-	x

No trabalho relatado em [HN88, LN98, NLH97], estes caracterizam-se por utilizarem uma abordagem hierárquica, partindo do princípio que a segmentação de baixo nível não é suficiente para separar informação relevante. Assim utilizam vários tipos de segmentação para gerar os dados de entrada do nível hierárquico superior.

Os trabalhos [ST10, GHS01] recorrem à utilização de MDS e MDT para encontrar a localização dos edifícios, evitando assim uma segmentação mais tradicional. O primeiro artigo determina os limites ajustando um polígono na localização do edifício enquanto

que o segundo opta por utilizar as técnicas de segmentação somente na localização do edifício.

O trabalho [SU08] utiliza uma segmentação baseada nas propriedades da cor do telhado ou da sombra do edifício. Os limites são gerados a partir de um vértice semente.

Na tabela 3.2 podemos consultar os dados de entrada destes artigos.

Tabela 3.2: Dados de entrada

Artigo	Fotografias	Densidade de informação	Modelos Digitais
[HN88]	nadiral	média	-
[LN98]	nadiral e oblíqua	média	-
[NLH97]	nadiral e oblíqua	média	-
[ST10]	nadiral IR	alta	nDSM
[GHS01]	satélite	média	nDSM e NDVI
[SU08]	nadiral	média	-

Iremos agora apresentar a análise de um conjunto de três artigos, que exploram uma abordagem interessante para identificar edifícios a partir de fotografias aéreas.

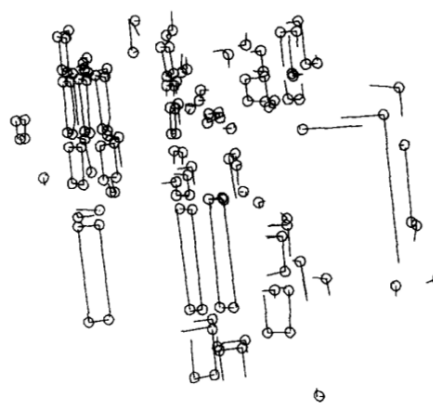
Estes utilizam fotografias de média densidade habitacional e um conjunto de técnicas de segmentação, em hierarquia, para irem refinando o resultado final.

A abordagem hierárquica advém do facto de que uma segmentação de baixo nível raramente proporcionar resultados completos e precisos[HN88].

A abordagem hierárquica utiliza um encadeamento de segmentações, na qual as de menor nível produzem os dados de entrada do nível seguinte. O método inicia-se com uma segmentação de arestas, a partir da qual é feita uma segmentação de linhas, figura 3.1(a) que por sua vez vai dar origem a uma segmentação de vértices, figura 3.1(b). As figuras anteriores são retiradas do artigo [HN88], que utiliza fotografia aérea do aeroporto internacional de Los Angeles. Na figura 3.2 é apresentado o mesmo encadeamento de segmentações, utilizado na dissertação[Car99], por outro autor.



(a) Segmentação de arestas



(b) Segmentação de vértices

Figura 3.1: Processamento de imagem

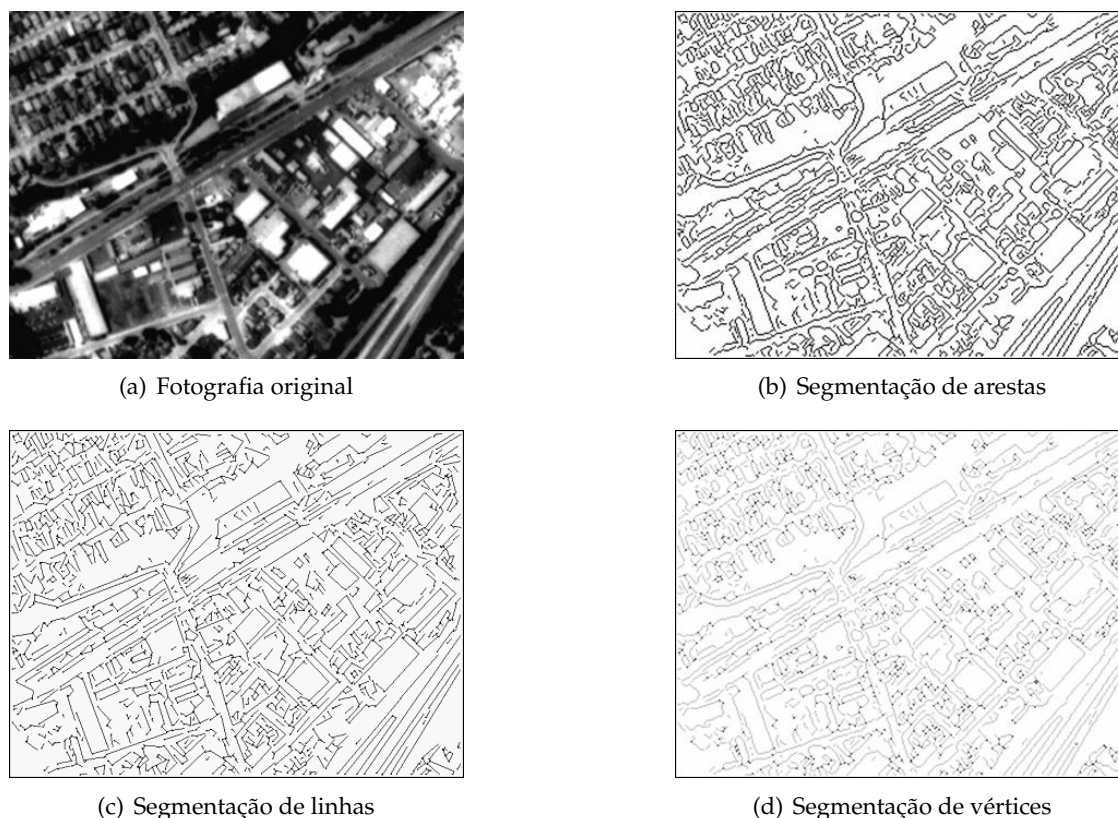


Figura 3.2: Processamento de imagem

Com a identificação dos vértices passa-se então a uma fase de classificação, sendo considerados vértices *bright* aqueles em que a média da intensidade da parte côncava for superior à da parte convexa, e *dark* os restantes.

Em termos de dados de entrada o artigo [HN88] utiliza fotografias aéreas nadirais e a informação da elevação do sol. Esta informação irá ser relevante em combinação com a segmentação de vértices, pois permite a identificação de sombras na fotografia.

Para efetuar a deteção, os edifícios, necessitam de cumprir um conjunto de características, como por exemplo, as paredes serem verticais e o telhado ser um plano horizontal.

Para "desenhar" os limites dos edifícios é utilizada a segmentação de linhas e as sombras. Com o resultado da segmentação de linhas é testada a continuidade da fronteira. De seguida é utilizada a localização da sombra, para inferir a presença de vértices ou arestas que não tenham sido detetados durante a segmentação. Na figura 3.3 podemos ver a relação entre a sombra e a fronteira, para efetuar a inferência dos vértices e arestas que se encontram a traço interrompido.

No entanto tal esquema apresenta um conjunto de limitações, nomeadamente no que se refere ao baixo contraste dos edifícios com o solo, o qual dificulta a segmentação feita nos níveis mais baixos. Por outro lado, e as sombras que são projetadas entre edifícios não permitem uma deteção adequada.

O segundo artigo [LN98] passa a utilizar como dados de entrada várias fotografias,

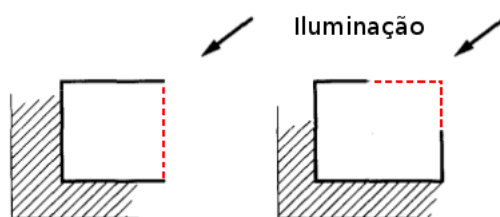


Figura 3.3: Utilização da sombra para inferir uma aresta ou um vértice

nadiraís ou de perspectiva oblíqua do mesmo local. As fotografias oblíquas possibilitam a utilização das paredes dos edifícios para contextualizar a informação. Este refere com maior pormenor as técnicas de segmentação utilizadas, inicializando o método com o algoritmo *Canny edge detector* e em seguida o *USC LINEAR* apresentado em [NB80].

Ao contrário do anterior, este identifica paralelepípedos, ver figura 3.4, que representam os edifícios. Assim, após a utilização das técnicas de segmentação são geradas, selecionadas e verificadas hipóteses, procedendo em seguida à análise e descrição 3D. Esta análise é feita através de evidências positivas como arestas, vértices, paralelismo ou sombras e de evidências negativas como a sobreposição ou as linhas cruzadas.

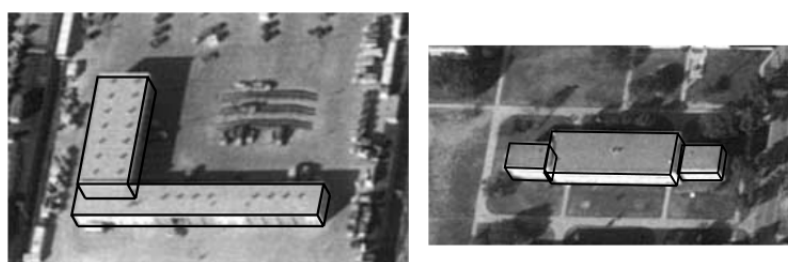


Figura 3.4: Identificação do edifício através de paralelepípedos

O artigo [LN98] introduz ainda o problema da diferenciação entre o fundo da fotografia e os objetos (*figure-ground problem*) e apresenta como limitações o elevado tempo de computação que é necessário para gerar os paralelepípedos.

O terceiro artigo [NLH97] apresenta a implementação da investigação feita nos artigos que se encontram anteriormente descritos [HN88, LN98].

Este adiciona a possibilidade de interação com o operador, que pode ser feita de duas formas, uma inicial e a outra corretiva. Na interação inicial o operador pode indicar, através de um cursor, a localização do edifício que pretende que seja detetado e o sistema irá ter em consideração todas as fotografias para gerar a melhor hipótese. Na corretiva o operador pode adicionar uma aresta que vai ser tomada em consideração na geração de hipóteses.

Enquanto que na interação inicial a hipótese para aquela localização é sempre gerada, na corretiva o sistema pode considerar que não existem indícios suficientes para gerar uma hipótese credível.

Este conjunto de três artigos abordam o problema da identificação de edifícios de uma

forma inovadora, pois executam um conjunto de segmentações para chegar ao resultado final. No entanto só estabelecem a altura do edifício, no caso desta dissertação pretendemos a cota do edifício. Seria necessário o auxílio de um MST ou um processo diferente para obter a altimetria destes objetos.

Iremos apresentar uma análise de dois artigos [GHS01, ST10] que utilizam modelos digitais como dados de entrada. Com a diferença entre o MDS e MDT é possível gerar o *normalized Digital Surface Model* (nDSM), que contém apenas a informação que se encontra acima do terreno. Na figura¹3.5(a) podemos ver os dois modelos MDS e MDT reais, enquanto que na figura² 3.5(b) está apresentado um esquema da operação realizada. Este modelo digital permite resolver o problema da diferenciação entre o fundo da fotografia e os objetos, que foi apresentado no artigo [LN98]. Apesar dos dois modelos digitais não estarem disponíveis para a realização deste trabalho, estes artigos apresentam uma forma diferenciada para o estabelecimento dos limites dos edifícios.

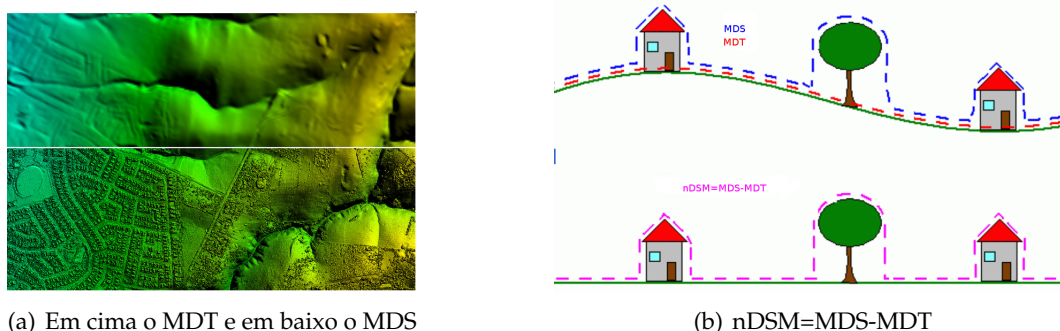


Figura 3.5: Modelos digitais

O artigo [GHS01] tem por objetivo a deteção de edifícios em ambiente urbano e utiliza fotografias de infravermelhos ortorretificadas. A extração dos limites é feita a 2D aplicando o conceito de polígono ortogonal que consiste, em enquadrar um polígono na área extraída. Para determinar o valor de altura, recorre ao nDSM para obter essa informação, ficando todos os edifícios com cobertura horizontal mesmo que esta não seja a realidade.

Além disso a solução proposta no artigo não tem em consideração as arestas nem os vértices, sendo que a inclusão destas características se encontra-se no trabalho a desenvolver.

O segundo artigo [ST10], tem por objetivo a extração automática de edifícios com cobertura retangular e circular em fotografias de satélite. Este adiciona ainda aos dados de entrada a informação da vegetação presente no *Normalized Difference Vegetation Index* (NDVI).

Além disso os autores recorreram à técnica de classificação *Support Vector Machine*

¹fonte: <http://aerometrex.com.au/blog/?p=89>

²adaptado: http://www.stadtentwicklung.berlin.de/umwelt/umweltatlas/ed610_03.htm

(SVM) [Han05] para obter classificação binária de áreas. Nas áreas de interesse é utilizado o algoritmo *Canny edge detector* para detetar as arestas e o algoritmo *Hough transform* [C62] para determinar os limites do edifício.

Apesar das funcionalidades oferecidas, a solução apresenta algumas dificuldades em detetar edifícios pequenos e que estejam próximos uns dos outros³.

O último artigo [SU08] utiliza fotografias de satélite e tem por objetivo a deteção de edifícios a 2D, mostrando-se relevante por utilizar técnicas de simples utilização. Para efetuar esta deteção utiliza somente técnicas de segmentação de cor, textura, arestas e a sombra, todas aplicadas no processo de determinação de regiões (*blob*).

É utilizada uma formula, associada à invariante da cor, e ao *threshold* de Otsu's [Ots79] para segmentar os telhados que tenham cor vermelha. Para o caso dos edifícios que não têm telhado vermelho, é usada uma variante desta fórmula para segmentar as sombras e a direção da luminosidade, determinando assim o centro do edifício.

Para construir os limites do edifício utiliza o algoritmo *Canny edge detector* para determinar arestas e em seguida é feita a identificação de vértices. Um destes vai ser selecionado como semente para iniciar a expansão do retângulo, dentro de uma margem de tolerância, até ser encontrado o vértice oposto. Podemos ver este processo na figura 3.6.

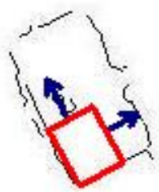


Figura 3.6: Expansão do retângulo

O artigo é inovador na forma como constrói os limites dos edifícios mas não deteta edifícios com formas não retangulares.

3.2 Extração baseada em métodos probabilísticos

Este método baseia-se na capacidade de detetar edifícios com o recurso a algoritmos de aprendizagem automática.

Os dois artigos [MLSB97, MLB⁺03] apresentados estão encadeados e pretendem demonstrar a capacidade deste tipo de algoritmos em resolver o problema da deteção de edifícios, na mesma fotografia e entre fotografias, para melhorar os sistemas de análise de imagem.

Como dados de entrada estes utilizam um conjunto de características que já se encontram disponíveis, não sendo necessário fazer essa análise e extração diretamente na fotografia.

³Em relação aos restantes edifícios presentes nas fotografias utilizadas nos artigos

Estes utilizam técnicas de aprendizagem supervisionada, com a classificação em duas classes, telhado (*rooftop*) ou terreno (*non-rooftop*) no caso do artigo [MLSB97], e em [MLB⁺03] é adicionada a classe desconhecido.

São utilizados nove tipos de valores todos eles contínuos, incluindo arestas, vértices, paralelismo, estruturas ortogonais de vértices, vértices junto a sombras, falhas de sobreposição, distância de arestas, linhas cruzadas e existência de junções.

Para o treino foram gerados 25 partições aleatórias do conjunto original, onde 60% foi utilizado para treino enquanto que os restantes 40% foram para teste.

No primeiro artigo [MLSB97] são analisados cinco algoritmos de aprendizagem, C4.5, CN2, *Nearest neighbor*, *Naive Bayes* e *Perceptron*.

Na fotografia utilizada foram geradas 1179 hipóteses para classificar, quando só existiam 29 edifícios. Devido ao grande número de hipóteses, foi implementado um programa para avaliar a região onde a hipótese se insere e fazer a classificação automática. Assim passam a ter 257 hipóteses para classificar pelo operador.

Os algoritmos foram executados de duas formas, treino e teste no mesmo conjunto de informação, mesma fotografia. E treino e teste feitos em conjuntos de informação diferentes, com o objetivo de determinar se seria possível treinar o algoritmo e depois aplicar em vários conjuntos de informação. Na primeira execução aquele que obteve melhor resultado a classificar os telhados foi o algoritmo *Naive Bayes*, mas também foi o pior a classificar os não telhados. Na segunda execução os melhores resultados foram do algoritmo CN2, conseguindo classificar corretamente todos os não telhados.

Estes testes demonstram, que na altura, não existia um algoritmo que pudesse ser utilizado com bons resultados em ambas as situações.

No segundo artigo [MLB⁺03] foram utilizadas fotografias de satélite e testados quatro algoritmos, designadamente o C5, *Nearest neighbor*, *Naive Bayes* e *Perceptron*.

Com a alteração de condições o algoritmo C5 comportou-se da mesma forma que o *Naive Bayes* no artigo anterior.

Uma das conclusões que o artigo revela é a necessidade de desenvolver uma forma de classificar a informação de treino, sem a interação de um operador.

3.3 Sombras

Uma sombra ocorre quando um objeto oculta total ou parcialmente uma fonte direta de luz. Esta pode ser dividida em sombra projetada ou própria, a figura⁴3.7 apresenta a diferença destes dois tipos de sombra.

A sombra própria é a parte do objeto que não se encontra diretamente iluminado pela fonte de luz.

Um objeto ao ocultar uma fonte de luz, projeta na superfície que se encontra no lado contrario, uma projeção da sua forma, denominada sombra projetada que por sua vez se divide em penumbra e umbra.

⁴adaptado de <http://sgvarts.blogspot.pt/2011/09/tip-solving-mystery-shadow-colors.html>

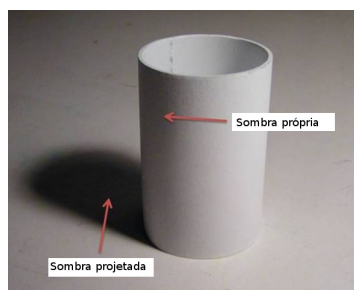


Figura 3.7: sombra projetada e própria

A parte da sombra que está a ser diretamente bloqueada pelo objeto denomina-se umbra, enquanto está parcialmente bloqueada é a penumbra[AGA08]. No entanto as fotografias que utilizamos nesta dissertação, não se identifica a presença da penumbra.

Existem aplicações em que estas características têm que ser tidas em conta, no que diz respeito ao resultado que se pretende desenvolver.

As sombras presentes nas fotografias aéreas, permitiram o desenvolvimento de técnicas com dois fins completamente diferenciados. Enquanto em certas aplicações se pretende utilizar este conhecimento para retirar informação, noutras o único objetivo é a sua eliminação.

Em termos de interpretação de imagem a presença de sombra, dificulta o trabalho dos operadores, devido à redução ou ausência de informação na imagem[Dar05].

Na elaboração de MDS a ausência de informação não permite a geração de pontos naquele local da superfície.

Este problema agrava-se quando estamos a trabalhar em áreas urbanas, pois a conjugação de edifícios de várias alturas, pode significar a perda de informação daqueles que se encontram na sombra dos mais elevados[Dar05].

Esta informação pode ser utilizada para determinar a altura de edifícios[Shi09, IM89, HN88], através da dimensão da sombra projetada e o valor angular da altura do sol, conforme podemos verificar na figura⁵ 3.8, Com a aplicação da equação 3.1.



Figura 3.8: Determinação da altura do edifício

⁵adaptado de <http://starsofeurope.weebly.com/solar-elevation-angle.html>

$$\text{Dimensão Edifício} = \text{Dimensão Sombra} * \tan(\text{Elevação do sol}) \quad (3.1)$$

Esta abordagem é utilizada em alguns dos artigos referidos, partindo de um conjunto de princípios que podem não estar presentes em todas as circunstâncias.

Um dos mais relevantes é o caso da sombra ter de ser projetada no terreno, como podemos verificar na figura⁶3.9, este pode trazer um conjunto de incorreções se o terreno adjacente não for horizontal.

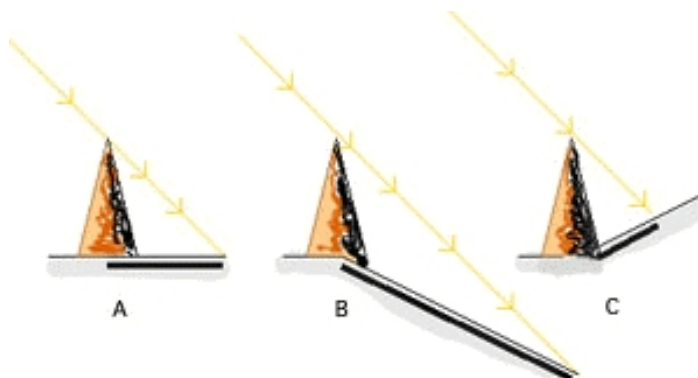


Figura 3.9: Diferenças de sombras devido ao terreno

Estes métodos de determinação de altura, acarretam um conjunto de incorreções que não são admissíveis para o trabalho em questão, mas permitem recolher um conjunto de ideias para determinar a existência de objetos construídos pelo Homem, pela presença de sombras.

O artigo [AGA08] tem por objetivo a deteção de sombras em fotografias de satélite de alta resolução. Não utiliza o ângulo da altura sol nem recorre a utilização de MDT.

Este método utiliza a invariante do espaço de cores e a informação de arestas para determinar corretamente as sombras.

Este encontra-se dividido em duas fases, na primeira são escolhidos pequenos conjuntos de píxeis que apresentam características de serem sombras. Estes são obtidos através do modelo de cores $c_1c_2c_3$ proposto no artigo [GS99]. Cada uma das componentes é determinada pela sua respetiva fórmula que tem por base uma transformação não linear do espaço de cor RGB.

Cada sombra é posteriormente caracterizada pela distribuição Gaussiana da componente c_3 dos valores dos píxeis dessa região.

Na segunda fase, recorrendo ao conjuntos de píxeis selecionados, são adicionados os píxeis adjacentes que respeitam a mesma distribuição. Para determinar com precisão o tamanho da sombra é utilizado um algoritmo de arestas para limitar a expansão da área.

O artigo [IM89] tem por objetivo a extração de edifícios de fotografias aéreas a partir da sua sombra. Uma das primeiras tarefas que é necessário realizar é a extração da área

⁶fonte: <http://www.vgl.org/webfiles/lan/cuspids/cuspids3.htm>

da sombra, que se encontra separada do restante método, para este efeito são utilizadas as seguintes técnicas, suavização da imagem e *thresholding*, fazendo a ligação destas regiões podem ser extraídas as zonas mais escuras como sombras.

Este artigo [Dar05] visa a deteção e remoção de sombras em fotografias de satélite monocrómicas, com aplicação em áreas urbanas de alta densidade. Este apresenta quatro algoritmos diferentes, mas potencialmente interligados, para separar regiões de sombra de regiões não sombra.

O algoritmo escolhido para realizar os testes na deteção de sombra foi o *thresholding* devido ao conteúdo do espectro, presente nas fotografias. Assim foi utilizado o *Bimodal histogram splitting*, que apenas apresenta duas elevações, entre as quais é feita a separação da informação da fotografia em sombra e não sombra.

3.4 Organização de cores

As cores, para serem visualizadas, processadas ou armazenadas, necessitam de ter uma representação formal e numérica. É através dos modelos de cores que estas são definidas em termos de representação matemática. A representação em componentes (eg: vermelho, verde, azul), permite a um equipamento interpretar, a cor, da mesma forma que o olho humano a interpreta, pelo brilho, saturação e a tonalidade[IHKM12].

Os modelos podem ser aditivos ou subtrativos. Nos modelos aditivos ao aumentar o valor aos canais, as cores tornam-se cada vez mais claras, resultando a combinação máxima de todos os canais no branco, na figura 3.10(a) está representada a conjugação de várias cores, o preto é representado pela ausência de cor e o branco por todas as cores. Os modelos subtrativos, o processo é praticamente inverso, ao acrescentar valor aos canais, as cores tornam-se cada vez mais escuras, na figura 3.10(b) está representada a conjugação das várias cores, o branco é representado pela ausência de cor, o preto por todas as cores[Wor].

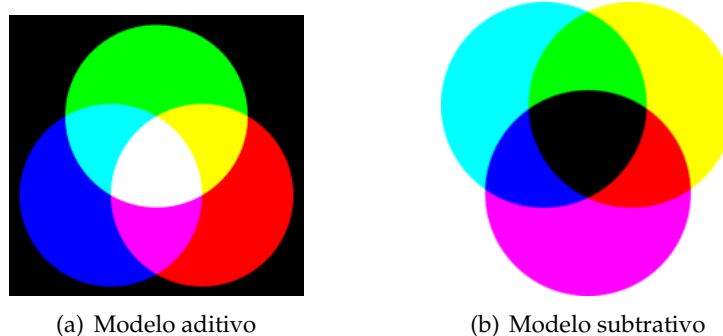


Figura 3.10: Modelos de cores

Outra forma de classificação, apresentada em[IHKM12], de acordo com as aplicações de processamento de imagem, cria as seguintes três categorias:

- Modelo de cores orientado ao dispositivo;

- Modelo de cores orientado ao utilizador;
- Modelo de cores independente do dispositivo.

Na Tabela 3.3 podemos visualizar esta classificação para alguns modelos de cores. Existe ainda uma tabela no artigo que apresenta as vantagens e desvantagens de cada um deles.

Tabela 3.3: Classificação segundo as aplicações do processamento de imagem

Modelo de cores	Classificação
Munsell	Independente do dispositivo
RGB, CMY(K)	Orientado ao dispositivo
YIQ, YUV, YCbCr	Orientado ao dispositivo
HSI, HSV, HSL	Orientado ao utilizador e dispositivo
CIE XYZ, CIE L*u*v*, CIE L*a*b*	Independente do dispositivo

O modelo de cores RGB, utiliza os canais vermelho, verde e azul, para efectuar a combinação de cores, podendo as coordenadas serem representada num cubo, onde o (0,0,0) significa o preto e (1,1,1) o branco se os valores forem normalizados para o intervalo [0,1], conforme se encontra na figura⁷ 3.11(a). Na diagonal deste cubo encontram-se os valores de cinzento. Este é considerado o modelo de cores base, pois não necessita de conversão para ser visualizado em monitores.

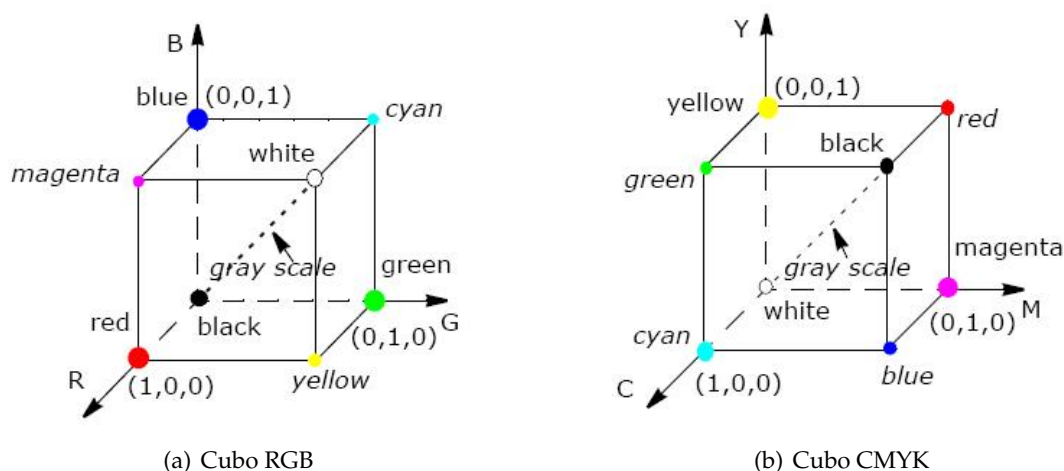


Figura 3.11: Representações de modelos

O CMYK é um modelo subtrativo, normalmente utilizado na impressão devido a utilizar as cores de impressão mais comuns (Ciano, Magenta, Amarelo e Preto). As coordenadas podem ser representadas num cubo, que se encontra representado na figura⁷ 3.11(b).

⁷ fonte: https://software.intel.com/sites/products/documentation/doclib/ipp_sa/80/ipp_manual/GUID-AE698C04-81DB-402B-88E7-2BEED820D4DF.htm

A família de modelos HSI, é composta pelo HSI, HSL e HSV, e encontra-se baseada no sistema visual humano, este utiliza coordenadas cilíndricas para a representação da cor. A diferença deste modelo prende-se com separação da intensidade(*chroma*) dos canais da tonalidade(H) e saturação(S), sendo que a intensidade está relacionada com a percepção humana. Em termos de representação, na figura⁷ 3.12(a), podemos visualizar o HSV que é representado por um cone, enquanto o HSI e HSL encontram-se representados na figura⁷ 3.12(b) por um duplo cone.

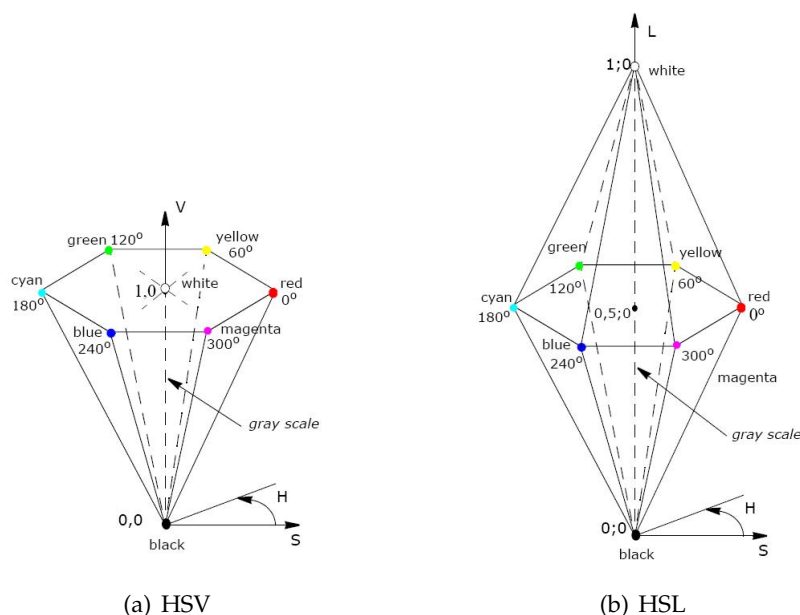


Figura 3.12: Representações dos espaços de cores

O modelo de cores CIE apresenta várias combinações, de espaço de cores, nomeadamente CIE XYZ, CIE $L^*u^*v^*$ e CIE $L^*a^*b^*$, que iremos de seguida apresentar.

O CIE XYZ só utiliza duas das coordenadas sendo o Z uma função das restantes duas. O CIE $L^*u^*v^*$, deriva do CIE XYZ, e referencia o ponto branco linearmente com o RGB(1,1,1). O canal L representa luminosidade, o u^* e v^* representam respetivamente a intensidade(*chroma*) e tonalidade(*hue*). No CIE $L^*a^*b^*$, os canais mantêm o mesmo significado, sendo que este foi construído com a intenção de ser mais perceptual em relação às alterações de cor ou seja uma alteração na cor é acompanhada por uma alteração da percepção[Poy96, p.17]. Em termos de representação este modelo utiliza vários planos, conforme se representa na figura⁸ 3.13(a), na figura⁹ 3.13(b) podemos visualizar o detalhe de um destes planos.

Os modelos da família YUV, são utilizados para efetuar a emissão de televisão, este inclui, o YUV, YIQ, YCbCr, YCC, YES. Este foram desenvolvidos para compatibilizar as emissões a cores com as a preto e branco, e efetuar a compressão da emissão. O canal Y representa a imagem em escala de cinzento enquanto que os restantes a cor. Desde que o

⁸ fonte: http://dba.med.sc.edu/price/irf/Adobe_tg/models/ciexyz.html

⁹ fonte: http://en.wikipedia.org/wiki/CIE_1931_color_space

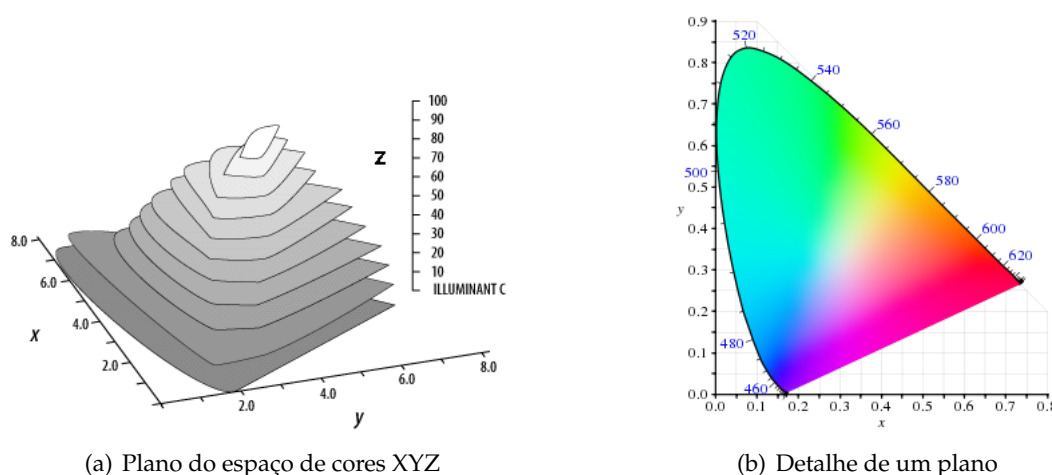


Figura 3.13: Planos do modelo CIE

canal Y chegue ao destino é possível visualizar a emissão.

No artigo [PV00] podem ser consultadas as informações necessárias para fazer as várias conversões entre espaços de cores. Existe por vezes a necessidade de fazer diversas conversões, quer seja de forma directa ou indirecta, passando por espaços intermédios, para chegarmos ao modelo ou espaço de cores que pretendemos.

O espaço de cores apresentado no artigo [CGZ08], é aqui apresentado devido a possuir características relevantes para a sua utilização neste trabalho. Como motivação, parte dos princípios da percepção visual com uma métrica de proximidade associada que sejam invariantes às alterações de luminosidade. Este espaço de cores pretende colmatar as necessidades de segmentação e *Poisson Image editing*[PGB03] em termos de processamento de imagem. Para fazer esta conversão é necessário partir do espaço de cores XYZ, aplicando em seguida duas matrizes apresentadas no artigo, que segundo as suas conclusões proporcionam melhores resultados, em termos de percepção, que os espaços de cores CIE $L^*a^*b^*$ e CIE $L^*u^*v^*$.

3.5 Detecção de pontos caraterísticos

As técnicas de detecção de caraterísticas permitem extrair um conjunto de pontos, cujas caraterísticas podem ser posteriormente utilizadas para identificar imagens semelhantes ou objetos que estejam contidos nas imagens. No decorrer deste trabalho houve a necessidade de estudar com maior pormenor duas destas técnicas, o *Scale-invariant feature transform* (SIFT) e o *Speeded Up Robust Features* (SURF). Estas foram seleccionadas pelo facto de serem diversas vezes referenciadas na bibliografia.

Ambos as técnicas se dividem em duas partes, na primeira são detetados um conjunto de pontos com caraterísticas relevantes, e uma segunda parte onde são criados os descritores a eles associados.

Existem outras técnicas para detetar caraterísticas de pontos que não são abordados

neste trabalho como por exemplo:

- Histogramas da orientação de gradientes [DT05];
- *Gradient location and orientation histogram* [MS05];
- *Binary Robust Independent Elementary Features* [CLF10];
- *Rotation invariant, fast feature* [TCT⁺13].

3.5.1 *Scale-invariant feature transform*

Esta técnica foi apresentada em 2004 no artigo [Low04], tem por objetivo a extração de características distintas e de forma independente da escala, em imagens, para posterior comparação, estando dividido em quatro fases:

1. *Scale-space Extrema Detection*: Efetua uma procura em diferentes escalas e localizações da imagem, utilizando um filtro *Difference-of-Gaussian* (DoG) para identificar potenciais pontos de interesse que sejam invariantes a alterações de escala e orientação.
2. *Keypoint Localization*: Em cada localização é adaptado um modelo para determinar a localização e escala. Os *keypoints* são escolhidos mediante a sua estabilidade;
3. *Orientation Assignment*: Com base no gradiente da imagem é atribuído a cada ponto uma ou mais orientações. Após esta informação ter sido extraída, todas as operações são realizadas sobre esta transformação, na figura 3.14(a) encontram-se desenhados os *Keypoints* e as suas respetivas orientações;
4. *Keypoint descriptor*: Transforma a informação numa representação que permita a sua utilização.

Esta técnica caracteriza-se por ser invariante a alterações de escala e rotação, permitindo a correspondência, de pontos, com adição de ruído, distorção afim, alteração da perspectiva 3D e alguma alteração de iluminação.

Nesta dissertação, esta será uma das técnicas utilizada para encontrar em ambas as fotografias, do modelo estereoscópico, para posterior correlação pontos notáveis.

3.5.2 *Speeded Up Robust Features*

Esta técnica foi apresentada em 2006 no artigo [BETVG08], e apesar de existir uma implementação oficial, esta encontra-se protegida por uma patente para efeitos comerciais.

Apresenta uma nova forma de solucionar o problema da deteção de pontos, em imagens, com escala e rotação diferentes, distinguindo-se dos restantes pela sua velocidade.

No detetor é feita uma aproximação através da matriz *Hessian* [HZ03, p. 598] devido à sua estabilidade e reprodução. Para cada ponto é calculada uma matriz, recorrendo ao operador de *Laplace* [Cyg07, p. 120].

Na figura 3.14(b), são apresentados os *Keypoints* de uma fotografia, utilizando a técnica SURF. Na figura 3.14 podemos efetuar uma comparação, entre os pontos característicos, detetados pela técnica SIFT e SURF.

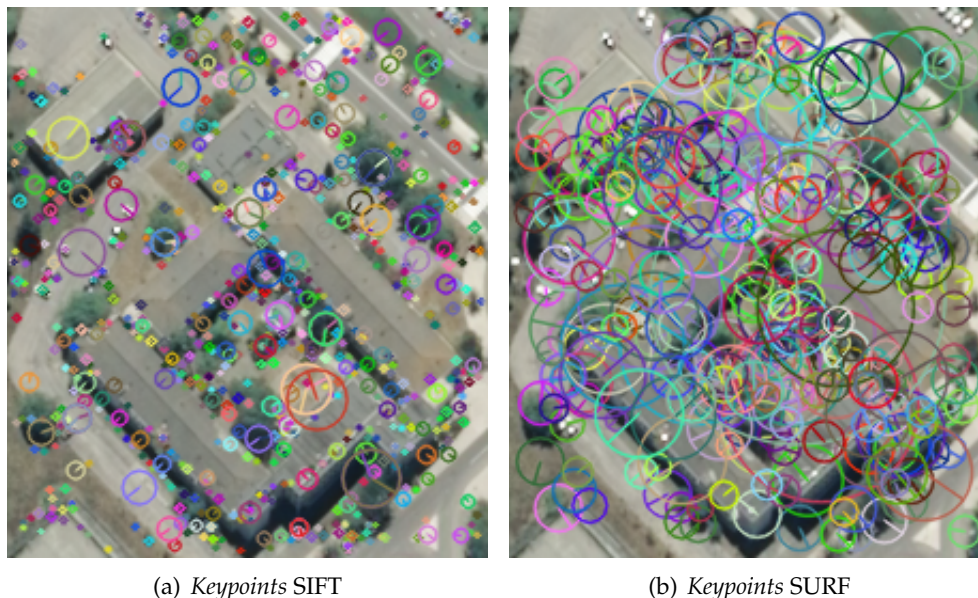


Figura 3.14: Conjuntos de *Keypoints*

Para produzir os descritores é utilizada a transformada de *Haar* [Sal06, p. 554] na proximidade do ponto, para as direções de x e y .

Esta técnica disponibiliza ainda a possibilidade de criar descritores, que não têm em consideração a invariante em rotação, que pode ser útil em casos específicos¹⁰, e que tornam o processo mais rápido.

3.6 Correlação de pontos

As técnicas de correlação de pontos analisados nesta secção, visam o estabelecimento de pontos homólogos entre fotografias.

Para obter estes resultados existem, basicamente duas abordagens, a primeira é a de comparar todos os pontos entre si. Este processo tem uma complexidade de $O(n \times m)$, e de forma geral é pouco atrativo devido ao tempo de computação necessário. A segunda abordagem é feita através da vizinhança (*nearest neighbor*), determinando com uma complexidade inferior os pontos homólogos, entre as duas fotografias.

Existem diferentes métodos para determinar pontos homólogos pela sua vizinhança, na realização deste trabalho utilizamos na técnica apresentada em 2009, no artigo [ML09]. Este não introduz nenhum algoritmo novo, proporcionando a utilização de dois algoritmos, numa mesma implementação, adaptando-se ao melhor algoritmo, mediante os dados de entrada.

¹⁰ Aplicações onde todas as fotografias são tiradas no mesmo plano

No artigo anteriormente referido são utilizados o algoritmo *randomized kd-tree* apresentado no artigo[SAH08] e o *Hierarchical K-Means tree* apresentado no artigo [LMS06]. Este menciona ainda que estes são dois dos algoritmos possíveis e testados, deixando em aberto a possibilidade de utilizar outros, que venham a surgir com melhores desempenhos.

A utilização desta técnica, esconde do utilizador qual dos algoritmos é usado, escolhendo o algoritmo através de validação cruzada (*cross-validation*) para identificar qual o melhor algoritmo a utilizar.

Esta escolha entre algoritmos, é necessária devido aos testes, realizados pelos autores, terem demonstrado que o algoritmo "ótimo" depende de vários fatores, como a estrutura dos dados de entrada ou a precisão pretendida.

Esta técnica é utilizada para efetuar a correlação de pontos, identificados pelas técnicas SIFT e SURF, conforme pode ser consultado na secção 4.3.2.1, onde podem ser visualizados exemplos da utilização desta técnica.

4

Trabalho Realizado

Genericamente pretendemos utilizar uma fotografia origem, do modelo estereoscópico, para obter a cobertura dos edifícios, através da informação radiométrica, eliminando os restantes objetos que não necessitamos de processar.

Com as coberturas dos edifícios identificadas na foto, pretendemos extrair os vértices dos polígonos que melhor representam a área da cobertura dos edifícios.

Para finalizar queremos encontrar os vértices homólogos, na fotografia de destino, sendo assim possível aplicar as equações de colinearidade para determinar as coordenadas 3D dos vértices que representam os polígonos. Para a implementação desta dissertação optámos por utilizar a linguagem de programação **C++**. Esta caracteriza-se por ser uma linguagem multi-paradigma, de nível médio. Na manipulação de imagens e aplicação de algoritmos de *computer vision*, utilizamos a biblioteca **OpenCV**, que possui uma interface para C++ e disponibiliza um vasto conjunto de funções. Esta é uma biblioteca *open source* com suporte para os sistemas operativos Windows, Linux e Mac OS.

Iremos agora descrever de uma forma sucinta o objetivo de algumas das **bibliotecas** implementadas no trabalho:

- **Globals**: armazena e permite a utilização de um conjunto de variáveis globais para parametrizar funções e métodos;
- **Utils**: possui um conjunto de funções que são transversais a várias funções e métodos, de utilização não específica;
- **UtilsSinglePhoto**: biblioteca que agrega todas as funções necessárias para manipular a informação de uma fotografia;

- *UtilsStereoPhoto*: biblioteca que agrega todas as funções necessárias para manipular a informação de um modelo estereoscópico;
- *Read e Write* : bibliotecas que permitem a leitura e escrita, em vários formatos de ficheiros.

Vamos agora passar a descrever a funcionalidade de cada uma das classes implementadas:

- *Collinearity*: classe que permite o cálculo da interseção direta espacial a partir das equações de colinearidade, de dois pontos homólogos, do modelo estereoscópico, e da aplicação do método dos mínimos quadrados;
- *Contour*: classe que permite guardar toda a informação de um contorno e restantes características¹, assim com aplicar um conjunto de funções para extrair informação do mesmo;
- *LineType*: classe para armazenar a informação necessária ao processamento de uma linha;
- *CornerType*: classe para armazenar a informação necessária ao processamento de um vértice;

Conforme foi referido na secção 1.3 o processo automático insere-se entre os processos de aerotriangulação e restituição. Na figura 4.1 é representada a constituição deste processo. O mesmo tem como dados de entrada as duas fotografias e os parâmetros de orientação externa do modelo estereoscópico. O resultado final é um ficheiro dos polígonos 3D das coberturas dos edifícios.

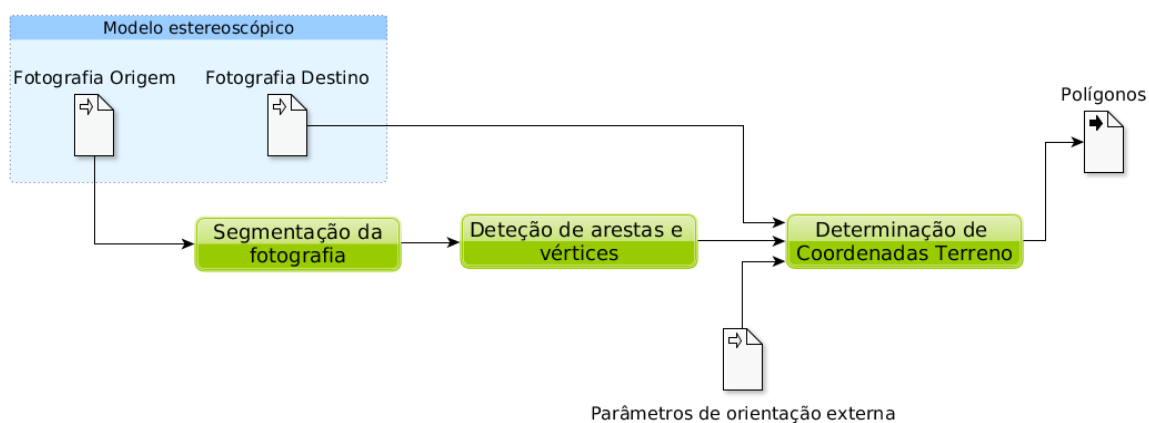


Figura 4.1: Componentes dos processo automático

¹Esta classe permite armazenar a informação de arestas e vértices, resultantes do processamento de um contorno. Assim como a informação de seu origem às arestas e vértices.

A implementação encontra-se dividida em três fases, estando assinaladas a verde na figura 4.1. Optámos por efetuar a separação para permitir uma divisão entre os diversos sub produtos e tornar o método mais modular.

Na primeira fase pretendemos efetuar a segmentação, da fotografia origem, recorrendo à informação radiométrica. Este processo visa a separação dos edifícios, dos restantes elementos presentes na fotografia, através da classificação e seleção das características radiométricas dos edifícios. Nesta fase é produzida uma imagem binária com a classificação dos píxeis pertencentes a edifícios.

Na segunda fase pretendemos efetuar a transformação das áreas, classificadas como edifícios, em polígonos a 2D.

Na terceira fase pretendemos obter as coordenadas 3D dos edifícios, através da identificação dos vértices 2D dos polígonos na fotografia de destino, e aplicando as equações de colinearidade, para determinar a coordenada 3D, passando de coordenadas foto para coordenadas terreno.

4.1 Segmentação da fotografia

Esta fase inicia-se com a conversão do espaço de cores, como foi abordado na secção 3.4. Devido a não se encontrarem disponíveis no OpenCV as funcionalidades necessárias, houve a necessidade de efetuar a implementação de algumas conversões.

Em termos de segmentação, foram selecionados o algoritmo MeanShift [FH06] e o processo de quantização, com recurso ao algoritmo de K-Means. Ambos os algoritmos efetuam uma análise de agrupamento (*clustering*). O OpenCV disponibiliza implementação para ambos e estes são mencionados em diversas publicações sobre a segmentação de imagem. Sendo um dos exemplos [BK08a, p. 298, pp. 479-483].

Todas as interações do operador localizam-se nesta fase, tendo este a possibilidade de indicar o número de cores que pretende na segmentação e selecionar a cor pela qual pretende analisar os resultados da segmentação. Como dados de saída, esta fase proporciona uma imagem binária, em que os píxeis de cor branca representam as áreas que processamos para identificar os edifícios e os de cor preta as áreas que não pretendemos processar, criando desta forma uma máscara.

Esta secção inclui ainda uma abordagem à técnica da criação de mapas de disparidades. Esta não é utilizada no método apresentado, sendo a sua relevância dada pela possibilidade de ser integrada com um processo de segmentação.

O esquema da figura 4.2 representa as etapas que são necessárias realizar nesta fase do trabalho, especificando o primeiro bloco apresentado na figura 4.1 do processo automático.



Figura 4.2: Esquema da segmentação da fotografia

4.1.1 MeanShift

Este algoritmo tem por objetivo encontrar máximos locais no conjunto de dados discretos. Tem a capacidade de ignorar *outliers*, utilizando *windows* para o processamento independente dos dados[BK08b, p. 337].

A implementação disponibilizada pelo OpenCV é baseada no artigo [CM99], que possui dois parâmetros:

- σ_s = resolução espacial;
- σ_r = resolução de distância.

Este algoritmo possui os parâmetros de entrada σ_r (vizinhança pela proximidade de cor) e σ_s (vizinhança pela proximidade espacial), são utilizados conjuntos de valores multi-dimensionais, nomeadamente os valores x, y, vermelho, verde e azul.

Para além destes dois parâmetros, definidos no artigo [CM99], a implementação permite ainda utilizar imagens em pirâmide². Efetuando a segmentação nos níveis mais elevados e à medida que estes diminuem faz o refinamento dos limites das regiões. Este parâmetro não foi aplicado, uma vez que provoca um efeito de *blur* em determinadas regiões. Na figura 4.3 podemos visualizar esse efeito com o valor de $\sigma_r = 20$ e $\sigma_s = 20$. A figura 4.3(a) não utiliza a pirâmides, enquanto a figura 4.3(b) utiliza o valor de cinco no parâmetro de pirâmides. Podemos verificar que na figura 4.3(b), existe um efeito de esbatimento entre a estrada e o campo agrícola, na parte superior da imagem, assim como nas zonas de vegetação.

Este algoritmo em termos visuais, aparentemente, parece funcionar relativamente bem, tornando a cobertura dos edifícios de cor única, figura 4.4(a). Analisando ao pormenor, o valor da cor no interior de uma cobertura, verifica-se que esta é composta por diversas cores, muito semelhantes. Na figura 4.4(b) foram preenchidas as áreas com maior número de píxeis tendo sido detetados cinco valores de cores.

²Na sua versão original este algoritmo, o parâmetro da pirâmide não existe.

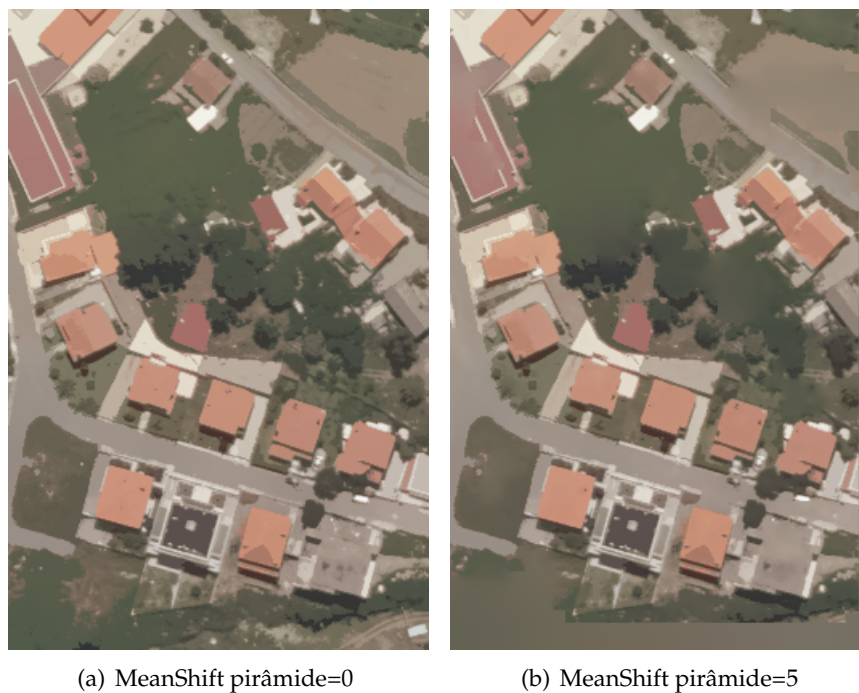


Figura 4.3: Algoritmo MeanShift com utilização de pirâmides

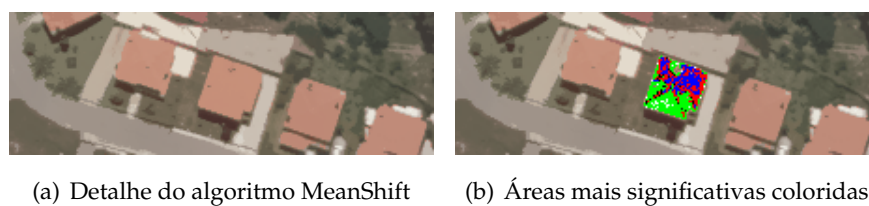


Figura 4.4: Algoritmo MeanShift

Na tabela 4.1 são descritas algumas das vantagens e desvantagens apresentadas por este algoritmo, utilizado na segmentação de fotografias.

Tabela 4.1: Análise ao algoritmo MeanShift

Vantagens
Permite a eliminação de informação
Proporciona resultados determinísticos
Desvantagens
Tem um custo computacional elevado para elaborar o resultado
Pode perde informação de arestas
Permite a existência de pequenas áreas dentro daquelas que pretendemos eliminar

4.1.2 Quantização do espaço de cor

O processo de quantização do espaço de cor, permite efetuar uma redução do número de cores presentes numa imagem, normalmente com o intuito de manter o mesmo aspeto visual. É sobretudo utilizado na adaptação de imagens, em termos de compressão ou para visualização em dispositivos com um número reduzido de cores.

O processo é aplicado em várias normas de compressão de dados, utilizando a *Discrete Cosine Transform* (DCT). Os exemplos mais significativos de normas que utilizam esta técnica são os ficheiros vídeo MPEG e imagens JPEG. A utilização da DCT permite bons rácios de compressão, mas não permite determinar o número de cores que queremos obter no final do processo.

Em concreto neste trabalho, o processo é utilizado para classificar as áreas, efetuando a segmentação da imagem. Assim utilizamos o algoritmo de aprendizagem não supervisionada K-Means[Han05], realizando uma análise de agrupamento *clustering*. Cada *cluster*³ é representado pela cor do seu centróide⁴, sendo o número de *clusters* um parâmetro definido pelo operador. Na figura 4.5 encontra-se um exemplo do processamento de uma fotografia, com diferentes valores no número de cores.



Figura 4.5: Processo de quantização

O algoritmo calcula *clusters* da informação radiométrica existente na fotografia. Retornando o centroide de cada *cluster*. Existindo a possibilidade da cor escolhida para representar o *cluster* não existir na fotografia.

Em termos gerais este funciona da seguinte forma :

1. Os parâmetros do algoritmo são o número de cores e as componentes do espaço de

³Grupo de valores que se encontram mais próximo do mesmo centroide.

⁴Valor que se encontram no centro de um grupo de valores.

cor de cada píxel;

2. Aleatoriamente são escolhidos píxeis cuja cor representará o centroide do *cluster*;
3. Associa cada píxel ao centroide mais próximo;
4. Recalcula o centroide com base na cor dos píxeis pertencentes ao *cluster*;
5. Volta para o ponto três até não haver alteração nos centroides⁵.

Este algoritmo tem a particularidade de convergir rapidamente, sendo o número de cores definido como parâmetro e calcula a proximidade com base na distância euclidiana[BK08b, p. 479], ao respetivo centroide.

O parâmetro que indica o número de cores é ajustado pelo operador, visto ser dependente da fotografia em questão. Nos testes que realizámos, tendo em consideração o segmento da fotografia com 262×400 píxeis, que representa 0.2 % de uma fotografia aérea, o valor de dez apresentou resultados aceitáveis, conforme se encontra representado na figura 4.5(b). Na figura 4.5(a) o valor de cinco é insuficiente, pois as coberturas dos edifícios ficam no mesmo *cluster* das estradas. Na figura 4.5(c), utilizando o valor de 100 como parâmetro as coberturas ficam muito segmentadas.

Na tabela 4.2 são descritas algumas das vantagens e desvantagens apresentadas por este processo, utilizado para segmentação.

Tabela 4.2: Análise ao algoritmo *Quantização*

Vantagens
Elimina a informação menos significativa
Necessita de pouca parametrização
Converge rapidamente
Desvantagens
Não proporciona resultados determinísticos
Necessita do número de <i>clusters</i>
Cálculo da proximidade entre píxeis

Este processo acaba por se encontrar mais adaptado ao objetivo de separar os edifícios dos restantes objetos, pois selecionando o número de cores, as coberturas dos edifícios, ficam representadas pela mesma cor, sendo possível fazer uma classificação.

O número de parâmetros necessários para a utilização deste processo são inferiores aos requeridos pelo MeanShift. A determinação do número de cores, requer menos preparação em relação à determinação do σ_s e σ_r , pois estes são de difícil visualização.

O tempo necessário para produção de resultados, é mais reduzido no caso do processo de quantização.

⁵Existe a possibilidade de definir um parâmetro, a partir do qual o movimento do centroide não seja significativo e o algoritmo termina.

4.1.3 Segmentação de arestas

Com o processo de segmentação concluído, é necessário contextualizar a informação, nomeadamente a do *cluster* que representa os edifícios.

Mediante a bibliografia consultada, o processo mais utilizado[GS99, LN98, ST10, SU08], passa pela aplicação do algoritmo *Canny edge detector*, apresentado no artigo [Can86], seleccionando as arestas de cada edifício. Este caracteriza-se por detetar arestas, com o mínimo de incerteza, em que os pontos pertencentes a cada aresta, estejam bem localizados ou sejam próximos da verdadeira representação da aresta.

Para as fotografias utilizadas neste trabalho, que se caracterizam por possuir uma elevada quantidade de informação, este algoritmo não proporcionou os melhores resultados.

Na figura 4.6 são apresentados alguns dos resultados obtidos com a utilização deste algoritmo. Ao analisarmos a figura 4.6(a), que possui valores de zero nos dois *thresholds*, podemos verificar que existe muito ruído⁶, nos dados de saída com estes parâmetros.

Na figura 4.6(c), ao inverter os parâmetros, colocando os *thresholds* em valores elevados, faz com que a informação dos edifícios seja perdida, mantendo mesmo assim algum ruído.

Ao efetuarmos o ajustamento manual dos parâmetros, eliminando o ruído, assistimos à perda de informação das arestas dos edifícios, como representado na figura 4.6(b).

Mediante as fotografias que temos disponíveis para esta dissertação, mesmo após o processo de quantização que reduz o número de cores, podemos considerar que o ajustamento dos *thresholds*, necessitam da intervenção de um operador com experiência, tendo sempre a presença de ruído no resultado final.

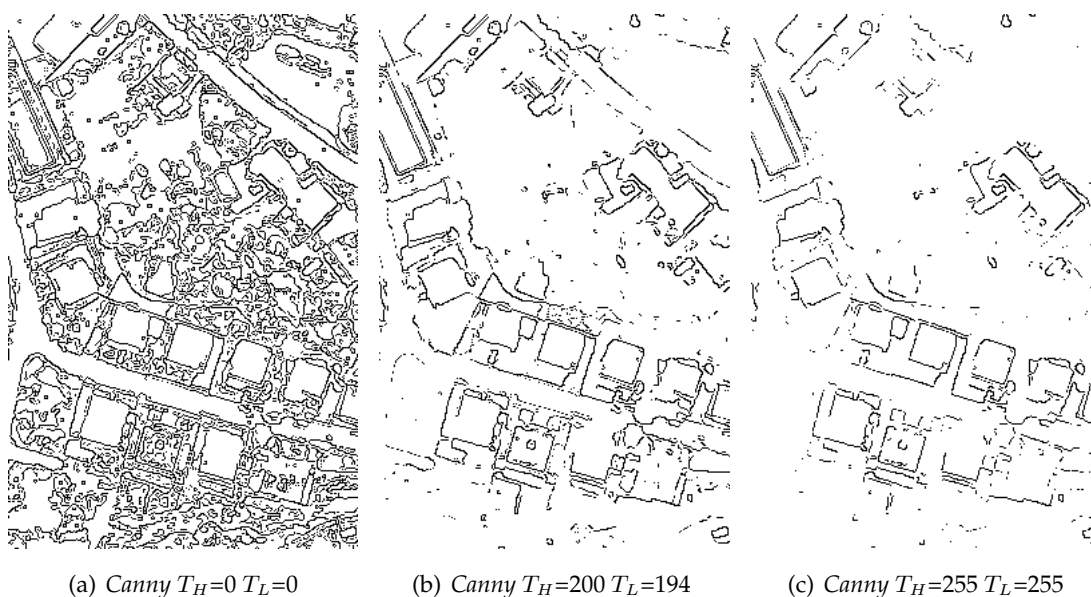


Figura 4.6: Algoritmo *Canny edge detector*

⁶Considerando que a informação, que não pertence a um edifício é ruído

Ao analisar os resultados obtidos podemos verificar que o facto de haver um píxel isolado, que não contém informação significativa, faz com que o algoritmo *Canny edge detector*, produza ruído, figura 4.7(a). O processo de quantização gerou um píxel isolado na cobertura de um edifício, ao aplicar o algoritmo *Canny edge detector* foram gerados oito píxeis em volta do píxel isolado. Desta forma é criado um ruído que teria que ser tratado posteriormente.

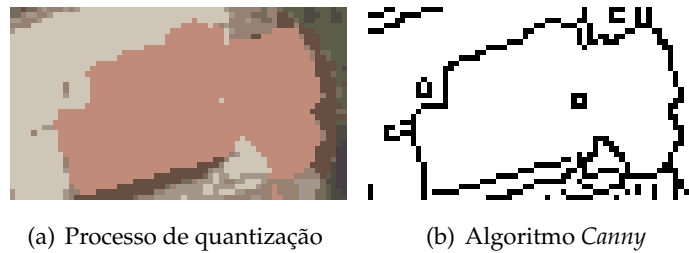


Figura 4.7: Resultado da aplicação do algoritmo *Canny* num píxel isolado

Outro aspeto que deve ser levado em conta é o facto de o algoritmo gerar uma localização o mais próximo possível. Esse facto leva a que o algoritmo *Canny edge detector* aplicado à figura 4.8(a), gere os píxeis representados na figura 4.8(b), fora dos objetos. Os objetos têm a mesma representação em ambas as figuras e encontram-se preenchidos a cinzento, enquanto a preto são representados os píxeis gerados a mais pelo algoritmo *Canny edge detector*.



Figura 4.8: Resultado da utilização do algoritmo *Canny edge detector*

Estes dois aspetos revelam que a utilização deste algoritmo gera píxeis em volta da cobertura, este efeito faz com que os edifícios sejam identificados por um contorno à sua volta, alterando assim a sua localização. O algoritmo [Can86] era um dos principais algoritmos que nos propusemos a utilizar, com a inviabilização do mesmo foi necessário seguir outra abordagem.

A nova abordagem inclui a interação com o operador. Este será responsável por determinar visualmente a região com a informação mais significativa dos edifícios. A partir

dessa região é criada uma máscara que vai ser utilizada para detetar características.

Outro método de definir a máscara pode passar por selecionar a região que tenha a cor mais próxima de um conjunto previamente estabelecido como cores frequentes na cobertura de edifícios. Este não foi utilizado, pois no território nacional não existe regularidade definida, globalmente, em termos das cores existentes de coberturas.

A forma mais *naïve* seria processar todas as máscaras. Este método torna o custo computacional extremamente elevado, pois obriga a processar o número de máscaras igual ao número de *clusters*.

De forma a simplificar o processo recorreremos ao operador para indicar quais as regiões que pretende processar. Ao efetuar um clique na fotografia segmentada, este indica a cor pela qual pretende que a máscara seja feita, sendo esta exibida com a opção de aceitar ou rejeitar.

O operador pode interagir aumentando ou diminuindo o número de cores que pretende na quantização, até obter a máscara ideal. Tem ainda a possibilidade de efetuar várias máscaras para identificar conjuntos de edifícios que tenham sido classificados em regiões diferentes. Na figura 4.9 é apresentada uma possível máscara.



Figura 4.9: Máscara

4.1.3.1 Quantização em vários modelos de cor

Conforme já foi referido a quantização tira partido do algoritmo de *clustering* K-Means, que gera *clusters* com base na proximidade das cores. Para efetuar o cálculo, da proximidade de dois píxeis, é determinada a distância entre ambos. O espaço de cores RGB apresenta um conjunto de limitações, relativamente ao cálculo da distância entre cores, em relação à percepção humana. Como diferentes espaços de cores possuem métricas diferenciadas, optamos por testar um conjunto de espaços de cores que apresentam uma melhor afinidade com o modo como o olho humano interpreta a informação.

Para obter resultados mais consistentes, utilizando este algoritmo, foram analisados os espaços de cor: HSV, HSI, YUV, CIELab e o espaço de cores apresentado no artigo [CGZ08], que designamos por pBased. O processo de geração destes resultados inclui:

- Leitura da fotografia no modelo original;
- Conversão para o modelo de cor;
- Aplicação do processo de quantização;
- Conversão da fotografia quantizada para o modelo original;
- Análise do resultado obtido.

Na figura 4.10 são apresentados os resultados da aplicação da quantização nos diferentes modelos de cor selecionados para a análise nesta secção. Para facilitar a visualização, as cores que se encontram na cobertura dos edifícios, foi alterada para destacar este conjunto de píxeis.

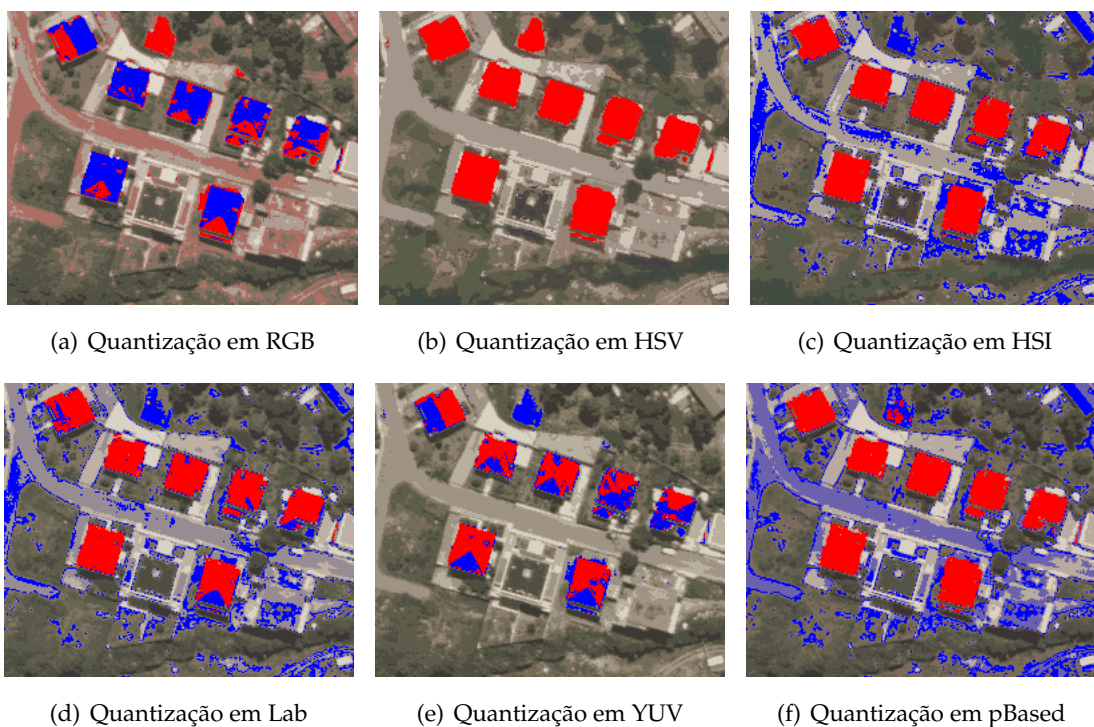


Figura 4.10: Processo de quantização com dez cores, em diferentes modelos de cor

Conforme podemos verificar o processo aplicado ao modelo RGB cria várias regiões na cobertura dos edifícios, conforme esperado.

Os modelos HSV e HSI que têm o canal *hue* separado, proporcionam resultados aceitáveis, sendo o HSV aquele que apresenta melhores resultados, uma vez que o HSI tende a esbater os limites dos edifícios.

O modelo Lab não gera resultados muito diferentes do modelo RGB.

O modelo YUV era o modelo em que, à partida, nos parecia oferecer maiores vantagens, devido à forma como armazena a informação. Contudo acaba por gerar resultados muito semelhantes aos do modelo RGB.

O modelo pBased, vocacionado para ser independente da intensidade da iluminação, apresenta resultados melhores que o modelo RGB, mas inferiores aos modelos mais simples como o HSV e HSI. Para efetuar o processo deste modelo é necessário realizar a quantização de todos os píxeis no formato de vírgula flutuante.

4.1.4 Mapa de disparidades

A geração de um mapa de disparidades permite determinar a distância que os objetos possuem em relação ao plano onde as fotografias foram tiradas. Com esta informação pretendíamos efetuar a separação, dos objetos elevados em relação ao terreno.

O mapa de disparidades de duas fotografias é calculado somente num dos eixos, sendo normalmente utilizado o eixo das abscissas.

A disparidade de pontos homólogos, é obtida através da distância entre as coordenadas de dois pontos, em duas imagens.

Para obter esse efeito é necessário efetuar a calibração e retificação/normalização das câmaras, ficando assim as linhas epipolares paralelas. Este processo garante que o píxel correspondente, na fotografia destino, se encontra na mesma linha do ponto que pretendemos.

O processo pode ser realizado píxel a píxel ou por bloco. O OpenCV permite a utilização de duas classes que efetuam a correlação de blocos. Nos testes realizados utilizamos a classe StereoSGBM, que aplica o algoritmo apresentado no artigo [HB05] de 2005.

Após ajustar manualmente os parâmetros é possível chegar a um mapa de disparidades conforme o apresentado na figura 4.11. Sendo que os píxeis mais claros se encontram mais próximos da câmara e os mais escuros o inverso.

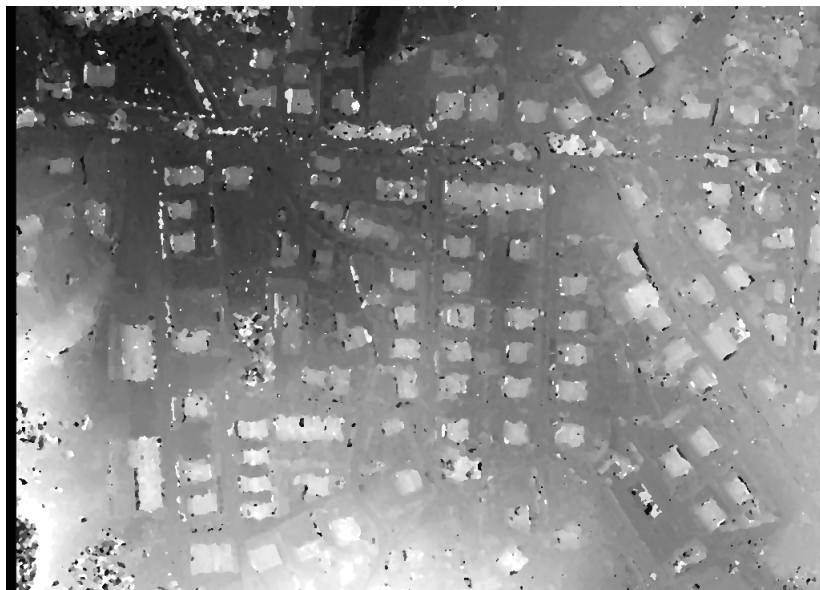


Figura 4.11: Mapa de disparidades

Com este processo obtemos uma representação da profundidade relativa à qual as superfícies se encontram, quanto mais próximo o objeto se encontrar da câmara mais clara é a sua representação.

Este processo aplicado a fotografias aéreas é limitado pelo efeito do terreno, conforme se pode verificar na parte inferior da figura 4.11, devido ao terreno ser mais elevado acabando por esbater a cor da cobertura dos edifícios, com a elevação do terreno.

O resultado do processo não permite a determinação de valores nas zonas que só existem numa das fotografias, denominadas oclusões. Estas são facilmente identificadas, por apresentarem zonas completamente brancas ou pretas.

Nesta dissertação, esta técnica não foi utilizada, pois implicava um estudo muito mais aprofundado da mesma, devido a existirem dificuldades em separar a profundidade dos edifícios do terreno, quando existem alterações na superfície terrestre.

4.2 Construção de polígonos

Na segunda fase é necessário transformar, de forma a simplificar, as diversas áreas obtidas. Para realizar esta transformação, detetamos um conjunto de características, que no final resultam num conjunto de polígonos em coordenadas foto, presentes na fotografia origem. Cada polígono representa a linha poligonal fechada da cobertura de um edifício. Para esse efeito efetuamos a deteção das seguintes características:

- Deteção do contorno de cada área da máscara;
- Deteção das arestas de cada contorno;
- Deteção dos vértices resultantes da interseção das arestas;

- Construção dos polígonos a partir dos vértices.

O esquema da figura 4.12 representa, de forma sucinta, as etapas que são necessárias realizar nesta fase do trabalho, especificando o segundo bloco apresentado na figura 4.1, do processo automático.

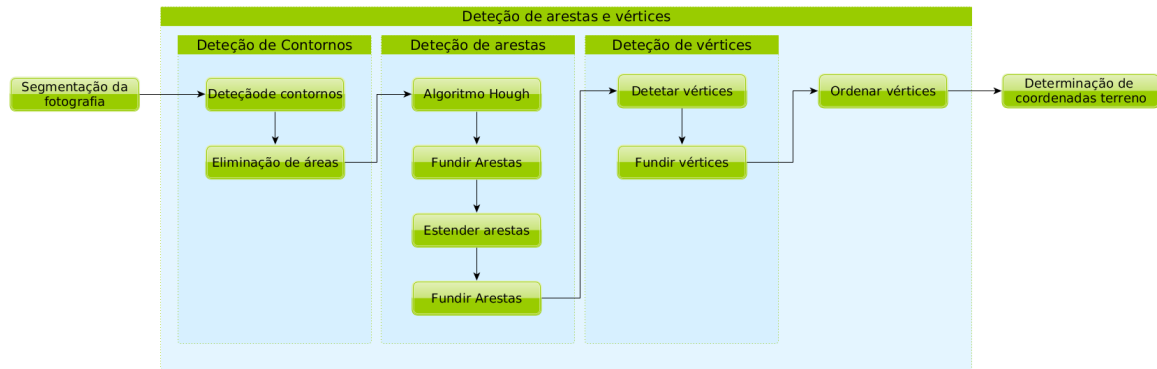


Figura 4.12: Esquema da detecção de arestas e vértices

4.2.1 Detecção de contornos

A detecção de contornos recorre a uma função disponibilizada pelo OpenCV, que implementa o algoritmo apresentado no artigo [SA85]. Este permite a detecção dos contornos, de uma imagem binária, fazendo o percurso que delimita toda a área.

A figura 4.13 representa o resultado da aplicação desta função a uma imagem contendo uma máscara 4.13(a) para a produção do respetivo contorno 4.13(b).

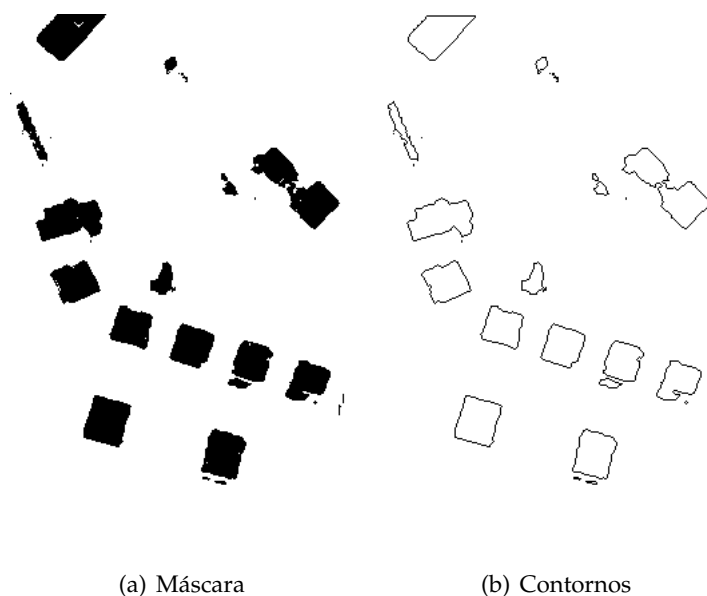


Figura 4.13: Resultado da função de detecção de contornos

Como dados de saída a função disponibiliza um vector, onde cada índice representa o contorno de uma área. Com a parametrização que utilizamos não é permitido à função gerar contornos interiores a outros contornos. Assim, não existe a necessidade de verificar se o contorno é externo ou interno a um edifício.

Cada contorno é constituído por um vector de objetos **Point**. A figura⁷ 4.14 representa o esquema deste tipo de estruturas.

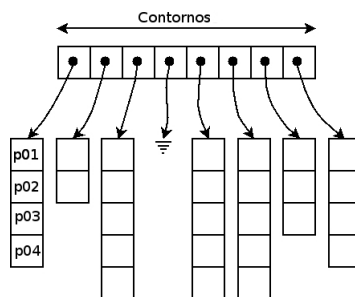


Figura 4.14: Vector de vectores de objetos *Point*

Esta forma de armazenar a informação, permite aceder aos contornos diretamente através do índice e, de seguida, iterar os pontos de cada contorno. Se quisermos visualizar o contorno, efetua-se a ligação de todos os seus pontos.

4.2.1.1 Eliminação de áreas

Na etapa anterior são produzidos contornos, que devido a serem de pequenas dimensões, não existe interesse no seu processamento.

Segundo as normas de aquisição do IGeoE, apresentadas na secção 2.3.1, não são adquiridos edifícios com menos de 3 m×3 m ou área equivalente. Para as fotografias aéreas usadas neste trabalho, uma área de 9 m² corresponde a um total de 36 píxeis, visto cada pixel corresponder a uma área de 0,52m×0,52m. Na pratica optamos por definir um *threshold* menor, de 25 píxeis, para evitar a geração de falsos negativos. Note-se que este valor poderá ser ajustado pelo utilizador. Efetuando uma comparação entre a figura 4.13(b), com todos os contornos, e a figura 4.15 que só contém contornos superiores ao valor determinado, verifica-se que foram eliminados doze contornos.

⁷Adaptado de <http://ycpcs.github.io/cs201-summer2014/notes/javaArrays.html>

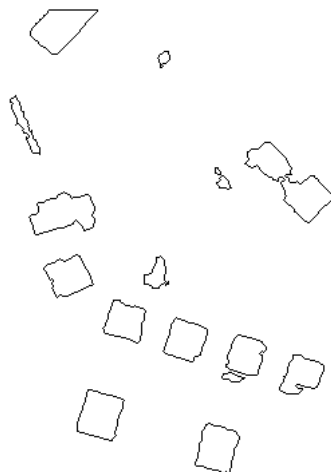


Figura 4.15: Eliminação de contornos

4.2.2 Detecção de arestas

A detecção de arestas tem por objetivo extrair, da informação disponibilizada pelo contorno, as arestas com maior significado na definição de um polígono. A importância desta etapa deve-se ao facto dos contornos não serem compostos por segmentos de reta, mas sim por um conjunto de píxeis que criam uma forma irregular. Assim sendo, existe a necessidade de efetuar uma transformação da representação raster para uma representação vetorial.

Inicialmente utilizamos o algoritmo probabilístico da transformada de Hough [MGK98] para detetar um conjunto de segmentos de reta compatíveis com pequenas secções de cada contorno. A aplicação deste algoritmo produz um elevado volume de segmentos. Para mitigar este problema, procedemos à eliminação de segmentos iguais, que tenham sido produzidos, sendo os restantes simplificados numa etapa de fusão.

A figura 4.16 representa os passos necessários para a detecção de arestas. Esta figura será utilizada para referenciar aspetos importantes na descrição de cada etapa.

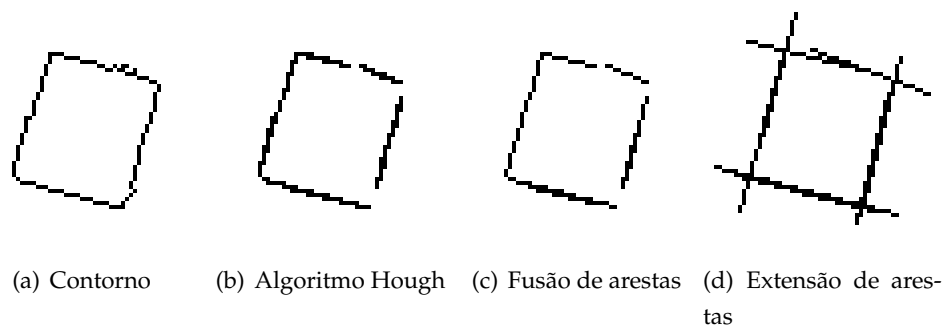


Figura 4.16: Arestas

4.2.2.1 Algoritmo de Hough

Para efetuar a detecção de arestas foi utilizado o algoritmo de Hough, com a seguinte parametrização⁸:

- O valor da distância ρ foi definido como 1, sendo a distância mínima do píxel;
- O valor do ângulo θ é variável entre os valores 0° e 90° ;
- O valor do *threshold*, foi definido como sendo 10, escolhido através de uma análise comparativa entre vários valores possíveis;
- O tamanho mínimo da linha foi definido como 6, pois o GSD é 0.52m logo são necessários seis píxeis para representar uma aresta de $\approx 3\text{m}$;
- O tamanho máximo do espaço entre píxeis da mesma linha foi definido como zero, pois os contornos são fechados.

O incremento da variação do ângulo θ , é parametrizável. Durante os testes apurámos que o valor de 0.01 é suficiente para detetar os segmentos de reta principais. A diminuição deste incremento aumenta o tempo de computação e o número de segmentos repetidos, sendo que o aumento dos restantes segmentos não é significativo. No Anexo A apresentamos uma breve análise para a determinação deste valor.

Na figura 4.16(b), encontra-se apresentado o resultado da aplicação do algoritmo de Hough, relativo à imagem apresentada na figura 4.16(a).

4.2.2.2 Fusão de arestas

Na etapa da fusão de arestas agrupamos os segmentos de acordo com a sua inclinação. Em cada grupo, caso os segmentos se intersetarem é criado um novo segmento que representa a orientação genérica do grupo. Esta etapa permite a redução do número de segmentos mantendo a orientação do grupo.

A figura 4.16(b) representa os segmentos de reta gerados pelo algoritmo de Hough, enquanto a figura 4.16(c) contém os segmentos fundidos. Apesar de visualmente estas duas imagens serem bastante semelhantes, a primeira foi construída desenhando 73 segmentos enquanto na segunda só foram utilizados 9. Foram necessários menos 88% da quantidade inicial de segmentos para representar a "mesma" informação.

Esta fusão é realizada conforme o algoritmo 1, sendo no final da etapa retornado um vector com os segmentos fundidos.

De referir que este algoritmo não tem a capacidade de fundir segmentos que sejam paralelos, devido a não existir interseção entre eles.

⁸Estes podem ser ajustados alterando o seu valor no ficheiro globals.cpp.

Algorithm 1 Merge all de lines in the array

```

1: function MERGELINES(linesVec, angle)
2:    $size \leftarrow linesVec.size.$ 
3:   for  $line \leftarrow 0$  to  $size$  do
4:      $l \leftarrow linesVec[line]$ 
5:     if  $l.processed \neq true$  then
6:       for  $lineSearch \leftarrow 0$  to  $size$  do
7:         if  $line \neq lineSearch$  then
8:            $ls \leftarrow linesVec[lineSearch]$ 
9:           if  $ls.processed \neq true$  then
10:             $minAngle \leftarrow ls.angle() - angle$ 
11:             $maxAngle \leftarrow ls.angle() + angle$ 
12:            if  $minAngle < l.angle < maxAngle$  then
13:               $interLines \leftarrow intersection(l, ls)$ 
14:              if  $interLines.processed = true$  then
15:                 $l \leftarrow interLines$ 
16:                 $ls.processed \leftarrow true$ 
17:              end if
18:            end if
19:          end if
20:        end if
21:      end for
22:    end if
23:  end for
24: end function

```

4.2.2.3 Extensão de arestas

Verificámos que os segmentos produzidos, na etapa anterior, não representavam a informação das arestas como pretendíamos, devido a dois fatores. O primeiro deve-se ao algoritmo de Hough criar vários segmentos numa mesma aresta, conforme se encontra representado nos dois segmentos da parte superior da figura 4.16(b). Este comportamento é expectável, pois não permitimos que o algoritmo de Hough gere segmentos que não sejam contínuos na imagem raster.

O segundo é devido aos contornos, na zonas dos cantos dos edifícios, possuírem uma ligeira curvatura. Este efeito não permite que os segmentos gerados terminem no final da aresta. Para solucionar estes problemas procedemos à extensão de todos os segmentos sendo estes, de seguida, novamente fundidos, conforme se encontra na figura 4.16(d).

Os segmentos de reta, em cada um dos lados, crescem mediante um parâmetro, previamente estabelecido, tendo sido utilizado o valor de seis neste caso particular.

4.2.3 Detecção de vértices

Para todos os pares de segmentos gerados na etapa anterior é determinado o ângulo e o respetivo ponto de interseção, formados pelas retas resultantes da extensão dos segmentos do par. Se o ângulo estiver compreendido entre os parâmetros de maior e menor

ângulo, aceites para identificar vértices, este é adicionado ao conjunto de vértices. Como os contornos dos edifícios são processados individualmente, vértices que se encontrem fora do retângulo criado para o contorno são eliminados. Quando é instanciada a classe *Contour*, o construtor gera os limites do retângulo onde o contorno se encontra inserido, com uma margem parametrizável. Este serve para efetuar a visualização das características detetadas e não permite "desenhar" fora deste. Desta forma a extensão de arestas e a deteção de vértices acaba por ser limitados por esta margem.

Na parte inferior da figura 4.17(a) podemos verificar que são gerados vértices fora da área do edifício. Efetuamos a eliminação dos mesmos tendo em atenção a distância que estes possuem para o contorno.

Na parte superior da figura 4.17(a) assistimos ao gerar de diversos vértices para representar o mesmo canto do edifício. Pois todos os pares de segmentos, geram vértices dentro do espaço criado para o contorno. Para simplificar esta informação efetuamos uma fusão de vértices utilizando a média ponderada do grupo de vértices.

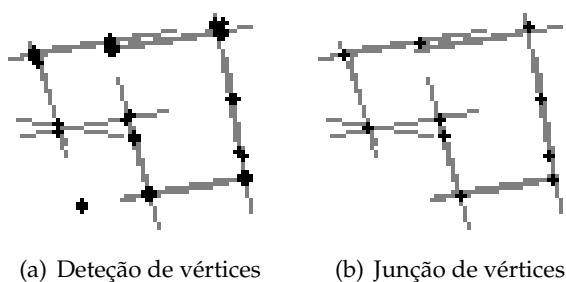


Figura 4.17: Vértices

A média ponderada tem em consideração, se a interseção ocorreu entre dois segmentos de reta, entre um segmento de reta e uma reta ou entre duas retas. Sendo que aquela que tem mais peso é a interseção entre dois segmentos.

Para definir se um vértice pertence ao grupo, é necessário ter em atenção que este não pode ter pertencido a outro grupo, não pode ter sido resultado de uma média ponderada já realizada e tem que ter uma distância mínima para o vértice que está a ser analisado, sendo esta parametrizável.

O pseudo código do Algoritmo que efetua a fusão dos vértices, é apresentado no algoritmo 2

No final da aplicação do algoritmo é retornado um vector com a localização média dos vértices.

4.2.4 Ordenação de vértices

Após as etapas de deteção de arestas e de vértices realizadas, temos um conjunto de pontos separados por polígono.

Algorithm 2 Merge all de corners in the array

```

1: function MERGECORNERS(cornersVec, minDistanceBetweenCorners)
2:    $size \leftarrow cornersVec.size.$ 
3:   for  $corner \leftarrow 0$  to  $size$  do
4:      $c \leftarrow cornersVec[corner]$ 
5:     if  $c.processed \neq true$  then
6:        $count \leftarrow 1$ 
7:        $tmpC \leftarrow c$ 
8:        $sumWeigth \leftarrow tmpC.getWeigth$ 
9:        $tmpC.setCorner(tmpC \times tmpC.getWeigth)$ 
10:      for  $cornerSearch \leftarrow 0$  to  $size$  do
11:        if  $corner \neq cornerSearch$  then
12:           $cs \leftarrow linesVec[cornerSearch]$ 
13:          if  $cs.processed \neq true$  then
14:             $distance \leftarrow calcDistance(c, cs)$ 
15:            if  $distance < minDistanceBetweenCorners$  then
16:               $sumWeigth \leftarrow sumWeigth + cs.getWeigth$ 
17:               $c.processed \leftarrow true$ 
18:               $cs.processed \leftarrow true$ 
19:               $count \leftarrow count + 1$ 
20:               $tmpC \leftarrow tmpC + (cs \times cs.getWeigth)$ 
21:               $tmpC.setWeigth(MAX(tmpC.getWeigth, cs.getWeigth))$ 
22:            end if
23:          end if
24:        end if
25:      end for
26:       $tmpC \leftarrow tmpC / sumWeigth$ 
27:       $mergeCornersVec.add(tmpC)$ 
28:    end if
29:  end for
30:  return  $mergeCornersVec$ 
31: end function

```

Os pontos de cada polígono, neste momento ainda não possuem qualquer tipo de ordenação. Sendo por isso possível, gerar diferentes polígonos, a partir do mesmo conjunto de pontos⁹.

Nesta etapa pretendemos encontrar o polígono que melhor representa um edifício. Assim os vértices têm que possuir uma ordenação que gere uma linha poligonal fechada para representar o edifício.

A ordenação dos vértices é feita pela ordem do ponto do contorno¹⁰, mais próximo do vértice. Assim, com os vértices ordenados, a representação do polígono é feita pela ligação consecutiva de pontos.

⁹Exceto se o numero de pontos for igual a três

¹⁰Num contorno todos os pontos possuem uma ordem.

4.2.5 Detecção de contornos (Convex hull)

O algoritmo de determinação do *Convex hull* permite simplificar um contorno efetuando a sua transformação para um contorno convexo.

A biblioteca OpenCV disponibiliza uma função para transformar um contorno convexo. Na figura 4.18 são representadas as diferenças no processamento, mediante cada etapa, se os dados de entrada forem o contorno (lado esquerdo) ou o *Convex hull* (lado direito) do contorno. Em termos de resultado final esta opção gera polígonos diferentes.

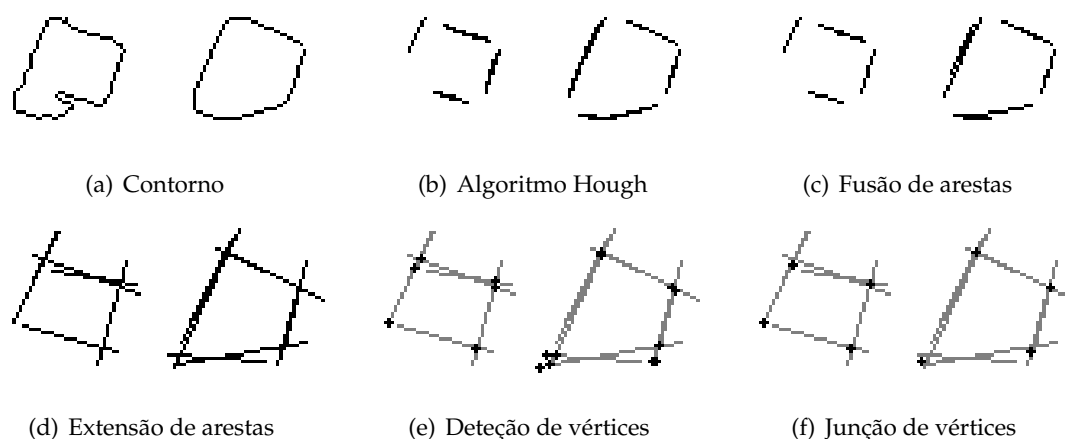


Figura 4.18: Processamento de contornos e *Convex hull*

De uma forma geral o processo é simplificado, mas os polígonos que não são convexos perdem a sua forma. Assim este algoritmo não foi utilizado.

4.3 Determinação de coordenadas terreno

O desenvolvimento das duas fases anteriores incidiu sobre a informação de uma fotografia, efetuando todo o processamento a duas dimensões e com coordenadas foto. Em termos de dados de saída, pretendemos agora passar essa informação para coordenadas terreno.

Para esta fase é necessário introduzir a informação radiométrica presente na fotografia destino, do modelo estereoscópico, assim como os parâmetros de orientação externa. Para determinar os vértices na fotografia de destino correspondentes aos da fotografia origem, são utilizados processos de determinação de pontos homólogos, evitando desta forma a necessidade de repetir as etapas realizadas na fotografia de origem para encontrar os vértices do polígono, além de que não haveria qualquer garantia de que fossem encontrados os mesmos vértices.

Na abordagem inicial pretendia-se utilizar métodos baseados na medição de ângulos e distâncias para encontrar os pontos homólogos, conforme será explicado na secção 4.3.6, sendo que estes não obtiveram resultados devidamente precisos para a realização deste trabalho.

Devido à não aplicabilidade da abordagem inicial, passamos a utilizar os princípios da geometria epipolar e de um coeficiente de correlação para determinar os vértices homólogos.

O esquema da figura 4.19 representa de forma sucinta as etapas que são necessárias realizar nesta fase do trabalho, especificando o terceiro bloco apresentado na figura 4.1, do processo automático.

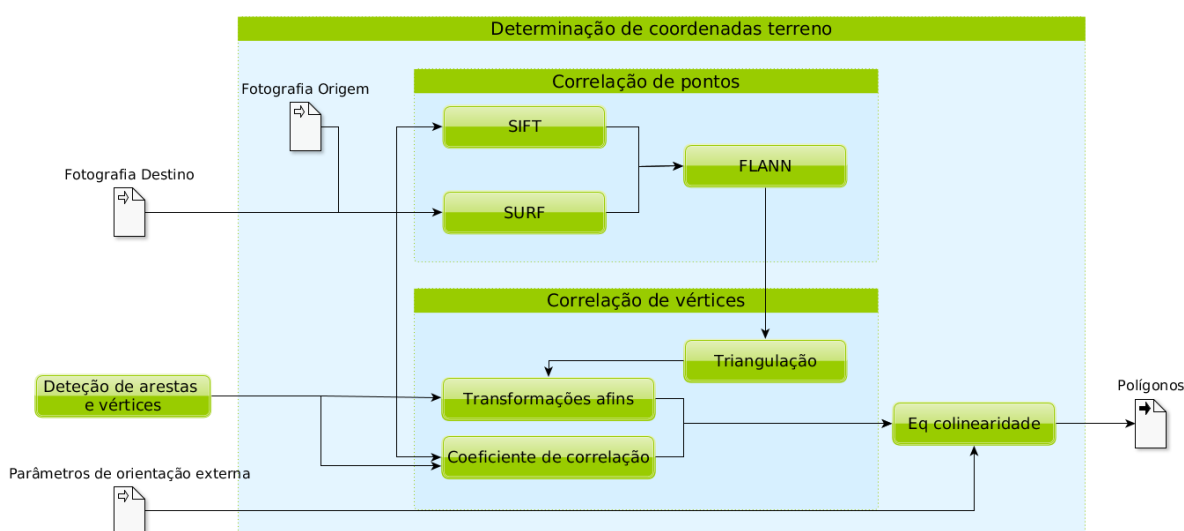


Figura 4.19: Esquema da determinação de coordenadas terreno

4.3.1 Transformações afins

A utilização de transformações afins, no âmbito deste trabalho tem o intuito de determinar vértices homólogos entre duas fotografias. Em geometria a utilização das transformações afins advém do facto de existir uma função entre espaços afins¹¹, que preserva a localização de pontos, segmentos de reta planos e o paralelismos entre segmentos. Mas não implica que seja possível preservar o ângulo entre segmentos de reta e a distância entre pontos [FPAW14].

Existe ainda a possibilidade de utilizar estas transformações em forma de matriz, tendo em atenção que para esta utilização é necessário utilizar coordenadas homogêneas. Desta forma a representação do vetor (x, y) será transformada em $(x, y, 1)$ para permitir a sua utilização na multiplicação de matrizes, conforme se encontra representado na equação 4.1. As coordenadas destino (x'_i, y'_i) são obtidas pela multiplicação da matriz afim com as coordenadas origem $(x_i, y_i, 1)$.

¹¹Uma estrutura geométrica que generaliza as propriedades afins de um espaço Euclidiano. Este não possui ponto de origem definido

$$\begin{bmatrix} x'_i \\ y'_i \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{bmatrix} * \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (4.1)$$

Para obter estas transformações, recorreremos a uma função, do OpenCv, para efetuar o calculo desta matriz afim. Para a sua utilização é somente necessário determinar um conjunto de pontos homólogos entre as duas fotografias, sendo que no mínimo são necessários três pontos.

Para a seleção dos três pontos recorreremos inicialmente aos pontos correlacionados da aerotriangulação. Estes eram selecionados pela menor distância ao que seria transformado e o ângulo interno seria superior a 50° entre si. A utilização dos pontos selecionados por este método gera vários problemas, pois nem sempre é possível encontrar três pontos que formem um triângulo com estas características.

Para simplificar a seleção dos pontos optou-se então por utilizar a triangulação de Delaunay, conforme vai ser explicado na secção 4.3.1.1. Esta técnica permite efetuar o pré-processamento de todos os triângulos e em seguida escolher aquele onde o ponto se encontra inserido, fazendo a verificação se o ponto se encontra no interior do triângulo.

Devido aos pontos da aerotriangulação estarem distanciados uns dos outros, se os triângulos gerados possuírem terreno acidentado, no seu interior, isso faria com que esta técnica apresentasse erros elevados.

Para colmatar a necessidade de pontos correlacionados, optou-se por correlacionar mais pontos para tornar esta técnica mais precisa. Foram então utilizadas as técnicas de correlação de pontos SIFT e SURF, que foram abordadas na secção 4.3.2.

A introdução de mais pontos permite criar triângulos mais pequenos e com um menor erros na determinação do ponto homólogo.

Apesar de esta nova técnica diminuir o erro entre os pontos, devido ao facto de estar a ser aplicada na superfície terrestre, apresenta de uma forma geral sempre um pequeno erro na localização do ponto.

Na tabela 4.3 é feita uma breve análise ao comportamento desta técnica.

Tabela 4.3: Análise da técnica da transformação afim

Vantagens
Baseada em princípios matemáticos e geométricos
Após a correlação não necessita da informação radiométrica das fotografias
Encontra sempre um ponto
Desvantagens
Modela o terreno como se estivesse definido por um plano
O ponto encontrado pode não ser o correto

4.3.1.1 Triangulação de Delaunay

Conforme já foi referido, a triangulação de Delaunay, foi utilizada para determinar os três pontos que se localizavam em volta do ponto que era necessário correlacionar.

Esta técnica matemática permite criar um conjunto de triângulos, no plano, em que os pontos de um triângulo são os únicos dentro de uma circunferência que passa pelos três, conforme está representado na figura¹² 4.20

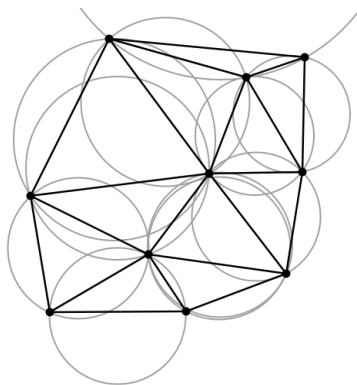


Figura 4.20: Triangulação de Delaunay e respectivas circunferências

A triangulação de Delaunay caracteriza-se por maximizar o menor ângulo de todos os triângulos gerados, evitando assim a existência de triângulos com ângulos internos pequenos, [BCKO08, pp. 196-205].

A biblioteca OpenCV disponibiliza uma classe que efetua todo o processamento da triangulação de Delaunay. Para a sua utilização é apenas necessário criar um objeto e inserir todos os pontos correlacionados nas duas fotografias. Após estas ações é possível aceder a um vector, onde cada índice representa um triângulo.

Cada triângulo encontra-se numa estrutura de dados, da biblioteca OpenCV, do tipo Vec6f. Esta estrutura guarda os valores x e y de cada um dos pontos.

4.3.2 Identificação de pontos homólogos

Para efetuar a identificação de pontos homólogos foram realizados testes com as técnicas descritas na secção 3.5. Estas permitem determinar os descritores de pontos, considerados distintos, em cada uma das fotografias.

Ambas as técnicas necessitam de classes diferenciadas para o detetor e extrator.

Para utilizar o SIFT é necessário definir os seguintes parâmetros no detetor:

- *contrastThreshold*: Serve para remover características em regiões de baixo contraste, quanto maior o valor menos características são produzidas pelo detetor;
- *edgeThreshold*: Utilizado para remover características das arestas, neste caso quanto maior o valor menos características são removidas.

¹²fonte: http://en.wikipedia.org/wiki/Delaunay_triangulation

Na classe do extrator é necessário definir o parâmetro *ampliação*, que determina o tamanho do descritor, fazendo a multiplicação da escala por este parâmetro.

Na utilização do SIFT foi utilizado o valor de 0.05 no parâmetro de *contrastThreshold* e 5.0 no *edgeThreshold* para o detetor e 1.0 no parâmetro de ampliação do extrator. Estes valores foram determinados através de testes onde foram apurados aqueles que melhor se adaptavam às fotografias disponibilizadas. De salientar que o valor do parâmetro de ampliação não implicou a alteração do número de pontos detetados e que o valor aqui apresentado é meramente exemplificativo¹³.

Após verificarmos que as duas técnicas eram válidas, optamos por utilizar os pontos detetados por ambas, aumentando assim o número de pontos disponíveis. Um exemplo dos *Keypoints* identificados foi apresentado na figura 3.14.

4.3.2.1 Correlação de pontos

Após termos um conjunto de pontos identificados em cada uma das fotografias, é necessário correlacionar os descritores, e identificar os pontos homólogos das fotografias. Utilizamos o *Fast Library for Approximate Nearest Neighbors* (FLANN), apresentado no artigo [ML09], e que já referimos na secção 3.6.

Em termos de implementação, a classe FLANN necessita somente, como parâmetros, dos descritores de ambas as fotografias. Os descritores da fotografia origem são utilizados para construir as estruturas de dados e denomina-se conjunto de treino. Os descritores da fotografia destino são utilizados para efetuar as consultas.

Como dados de saída esta classe disponibiliza um vector de objetos **DMatch**. Esta encapsula a informação:

- Distância entre descritores, do vizinho mais próximo;
- Índice da imagem de treino. Não é aplicável neste caso porque só utilizamos uma imagem para treino;
- Índice do descritor do conjunto de treino;
- Índice do descritor do conjunto de consulta.

Neste vector de resultados todos os pontos possuem um vizinho mais próximo. Este facto, isoladamente, não indica que estes sejam pontos homólogos, pois os descritores podem distar bastante uns dos outros. Na figura 4.21 é apresentado o resultado da detecção de pontos SIFT com a correlação FLANN. Como se pode verificar todos os descritores na fotografia de origem tem um vizinho mais próximo na fotografia de destino, mas só estamos interessados naqueles que são homólogos.

Para encontrar os pontos homólogos, foi utilizada a técnica apresentada no tutorial da página oficial do OpenCV¹⁴. Esta consiste em determinar o menor valor de todos

¹³Foram feitos testes com o valor de 1.0 e 300.0, no caso das fotografias utilizadas não houve alteração no número de pontos.

¹⁴http://docs.opencv.org/doc/tutorials/features2d/feature_flann_matcher/feature_flann_matcher.html



Figura 4.21: Correlação de pontos com a técnica FLANN

os vizinhos e em seguida só são considerados pontos homólogos, aqueles que têm um vizinho a duas vezes a distância mínima¹⁵.

Na figura 4.22 encontram-se os pontos homólogos encontrados através do FLANN e pela técnica descrita anteriormente. Na figura 4.22(a) foi utilizada a detecção de características SIFT. Comparando este resultado com o da figura 4.21, a diferença entre estes dados de saída encontra-se na eliminação dos pontos que não são homólogos. Na figura 4.22(b) é apresentado o mesmo processo com a utilização do detetor de pontos característicos SURF.



(a) SIFT e Pontos homólogos



(b) SURF e Pontos homólogos

Figura 4.22: Pontos homólogos

¹⁵O método apresentado pelo autores do artigo, implica o calculo do rácio da distância em o vizinho mais próximo e o segundo mais próximo.

4.3.3 Coeficiente de correlação

Com a utilização das transformações afim, encontramos os vértices homólogos, sem ter em atenção a informação radiométrica disponibilizada pela cor do píxel.

A utilização do coeficiente de correlação permite verificar a semelhança da informação tendo por base os seus dados. Para que seja utilizado é necessário estabelecer duas matrizes, a matriz padrão, centrada no píxel que queremos encontrar na fotografia de origem, e a matriz busca que é centrada no píxel que queremos avaliar na fotografia destino.

Existem diversas fórmulas para efetuar o cálculo deste coeficiente. Nesta dissertação utilizamos o processo do coeficiente de correlação de **Pearson** ou produto-momento, conforme apresentado na equação 4.2.

$$r = \frac{\sum_{i=1}^n (p_i - \bar{p})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^n (p_i - \bar{p})^2} \sqrt{\sum_{i=1}^n (b_i - \bar{b})^2}}, \quad -1 \leq r \leq 1$$

p – matriz padrão
 b – matriz busca
 \bar{p} ou \bar{b} valor médio de cada uma das matrizes

(4.2)

O valor de r encontra-se compreendido entre -1 e 1 sendo que:

- $r = -1$: A correlação negativa, as duas matrizes variam em sentidos opostos;
- $r = 0$: Não existe correlação;
- $r = 1$: A correlação é perfeita entre as duas matrizes;

Para este caso concreto estamos interessados nos valores compreendidos, entre 0 e 1, intervalo no qual as semelhanças, da informação radiométrica, das matrizes tendem a aumentar ou diminuir simultaneamente, na tabela 4.4 encontra-se apresentado uma classificação proposta em [San07].

Tabela 4.4: Classificação da correlação

Coeficiente de correlação	Correlação
$r=1$	Perfeita Positiva
$0.8 \leq r < 1$	Forte Positiva
$0.5 \leq r < 0.8$	Moderada Positiva
$0.1 \leq r < 0.5$	Fraca Positiva
$0 \leq r < 0.1$	Ínfima Positiva
0	Nula

O tamanho da matriz a utilizar, é parametrizável, e foi alvo de teste para determinar, aquele que melhor resultado apresentava, tendo em atenção que conforme maior for a matriz, maior será o tempo de processamento.

Devido à cor das fotografias estar decomposta em três canais, este coeficiente é determinado sobre os valores de cinzento. Existe a possibilidade de aplicar, o coeficiente, em cada um dos canais separadamente, mas esta técnica não foi utilizada neste trabalho.

Em termos da seleção dos pontos na fotografia de destino, aos quais deve ser determinado o coeficiente de correlação, recorreu-se os princípios de geometria epipolar, mais concretamente a restrição epipolar [Cyg07, pp.31-36]. Conforme se pode verificar na figura 4.23¹⁶, sabendo a translação e rotação, relativa das duas câmaras, e a localização do ponto X_L , a posição do ponto X_R , encontra-se na linha definida por $e_R - X_R$. Assim é possível verificar os pontos da linha e encontrar o ponto específico X_R através do coeficiente de correlação. Esta linha denomina-se linha epipolar.

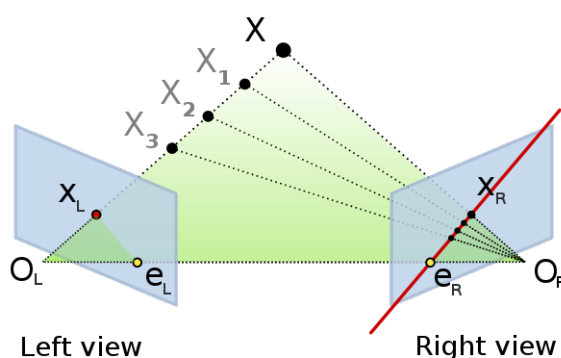


Figura 4.23: Restrição epipolar

Para determinar a linha epipolar é utilizada a matriz fundamental. Esta é uma representação algébrica da geometria epipolar, que permite o mapeamento entre um ponto e a sua respetiva linha epipolar [HZ03].

Para obter a matriz fundamental, que relaciona as fotografias que são utilizadas nesta dissertação, recorreremos ao OpenCV. Este permite o cálculo da matriz, fornecendo como dados de entrada um conjunto de pontos homólogos, pelas técnicas SIFT e SURF, e o método de cálculo. Para este trabalho optamos por utilizar o método *RANdom SAmple Consensus* que necessita no mínimo de oito pontos. Este método foi selecionado devido a ter sido concebido para a área científica da produção cartográfica, tendo sido apresentado no artigo [FB81]. Como dados de saída, é produzida uma matriz 3×3 , que conjuntamente com o valor do ponto, servem de dados de entrada para calcular o vector da linha epipolar.

O vector é transformado num segmento de reta, e em seguida criado um iterador de linha, para percorrer todos os pontos. Para cada píxel da linha é calculado o coeficiente de correlação, sendo o maior valor encontrado considerado o ponto homólogo.

A utilização desta técnica demonstrou ser bastante fiável, determinando quase sempre o píxel que pretendia-mos.

¹⁶fonte: http://en.wikipedia.org/wiki/Epipolar_geometry

O coeficiente de correlação pode ainda ser aplicado em fotografias retificadas/normalizadas, utilizando o processo que se encontra descrito na secção 4.3.6.3. Desta forma não é necessário determinar a linha epipolar, podendo a pesquisa ser realizada na linha horizontal da fotografia de destino. Neste caso como as imagens foram transformadas é necessário efetuar uma conversão do ponto para o novo sistema, fazer a pesquisa com o processo do coeficiente de correlação, e para finalizar, converter de volta o ponto encontrado para o sistema original da fotografia.

Este processo não foi realizado, pois era necessário efetuar o processamento adicional de retificação/normalização, que acarreta bastantes custos.

4.3.4 Conjugação de dois métodos

Depois de avaliarmos vários métodos para determinar pontos homólogos, houve necessidade de escolher aquele que mais se apropriava ao objetivo que pretendemos alcançar.

Os métodos implementados foram:

- Interseção direta¹⁷;
- Cevianas¹⁷;
- Transformações afins;
- Coeficiente de correlação;
 - Retificação/normalização de fotografias¹⁷;
 - Geometria epipolar;

Foram testados vários processos para efetuar a correlação de vértices, tendo sido os mais significativos os seguintes:

- O processo de transformações afins, que apresenta um erro sistemático na localização do vértice, mesmo que pequeno, garante que o ponto se encontra compreendido dentro de um triângulo;
- O processo do coeficiente de correlação que possui uma elevada taxa de êxito na localização do vértice, mas caracteriza-se por possuir um custo computacional elevado pois tem de analisar uma elevada quantidade de píxeis.

Com a combinação de ambos, podemos complementar as suas propriedades. Utilizamos as transformações afins para diminuir a área de pesquisa e o coeficiente de correlação para encontrar o vértice.

Foram efetuadas diversas aproximações para aferir qual seria a mais apropriada para este método.

Inicialmente, foi utilizado o coeficiente de correlação para fazer a pesquisa dentro do triângulo, conforme o exemplo da figura 4.24(a). Neste caso o triângulo pode conter um

¹⁷ Apresentado na secção 4.3.6

conjunto elevado de píxeis e não utiliza toda a informação disponibilizada pela transformação afim e restrição epipolar.

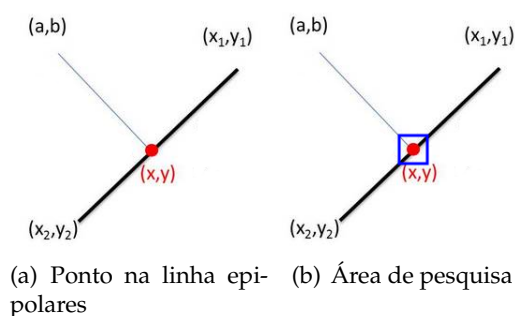


(a) Pesquisa no interior do triângulo (b) Pesquisa na linha epipolar (c) Pesquisa na região calculada

Figura 4.24: Região de pesquisa

Com o intuito de diminuir a área de pesquisa, a aproximação seguinte foi utilizar o segmento da linha epipolar que se encontra dentro do triângulo, conforme o exemplo da figura 4.24(b). O número de píxeis diminui significativamente mas continua a não utilizar toda a informação da transformação afim. Outro dos fatores para eliminar esta aproximação deve-se à possibilidade dos pontos homólogos utilizados não terem qualidade, fazendo com que a linha epipolar possa passar ligeiramente ao lado do ponto¹⁸.

A aproximação final foi a de aproveitar a informação total dos dois métodos. Inicialmente é determinado o ponto (x,y) da linha mais próximo do ponto (a,b) da transformação afim, conforme apresentado na figura¹⁹ 4.25(a), e ao redor deste novo ponto (x,y) criar-se uma área de pesquisa, figura 4.25(b).



(a) Ponto na linha epipolar (b) Área de pesquisa

Figura 4.25: Ponto da linha mais próximo a outro ponto

A área é calculada tendo em consideração a multiplicação da distância $\overline{(a,b),(x,y)}$, por um valor constante, que nos garanta uma taxa de êxito, mesmo se a linha epipolar esteja desviada. No caso desta distância ser zero, a distância considerada, para o efeito do cálculo da área, é de um, obrigando desta forma a existir sempre pesquisa.

¹⁸Acontece quando os pontos não estão bem distribuídos.

¹⁹adaptado de <http://stackoverflow.com/questions/5204619/how-to-find-the-point-on-an-edge-which-is-the-closest-point-to-another-point>

Existe ainda espaço para melhoramentos neste método, pois este acaba por dar mais importância ao método da linha epipolar do que ao da transformação. Existem várias hipóteses, que poderão ser testadas, para possíveis melhoramentos:

- Balancear a área de pesquisa para um dos lados da linha;
- Aumentar ou diminuir a área de pesquisa mediante a forma e área do próprio triângulo;
- Utilizar um círculo como área de pesquisa, em vez do quadrado.

4.3.5 Equações de colinearidade

Após terem sido realizadas as etapas da localização dos vértices homólogos, obtemos a localização do ponto na fotografia de origem e na fotografia de destino. Estas são coordenadas foto, bidimensionais, sendo necessário transformá-las em coordenadas terreno, tridimensionais.

Conforme já foi referido, na secção 2.4, vamos utilizar as equações de colinearidade. Para este efeito, foi implementada uma classe para encapsular todo o cálculo das equações e executar o processo do método dos mínimos quadrados.

Em termos de parâmetros de orientação interna, existe um construtor que carrega estes valores para o conjunto de fotografias do bloco que nos foi cedido e utilizado nesta dissertação, ou então existe a possibilidade de utilizar um construtor onde são definidos todos estes parâmetros.

As orientações externas, da cada fotografia, têm que ser inseridas através de um método próprio, com os parâmetros Ω , ϕ , κ e coordenada 3D do centro de projeção da fotografia.

Para iniciar o cálculo é necessário utilizar o método *compute*, com os valores dos pontos homólogos de cada fotografia, em milímetros, e a precisão pretendida no método dos mínimos quadrados. Em retorno o método devolve uma coordenada terreno.

A mesma classe pode calcular o número de pontos que for necessário, fazendo várias iterações com o método *compute*.

Para fazer a importação da informação gerada do método pelo ArcGIS, são criados dois ficheiros. Um no formato de texto e outro no formato KMZ, ambos com a informação estruturada em polígonos. Através da ligação existente entre o ArcGIS e a aplicação de restituição Summit Evolution, os polígonos gerados são visualizados a 3D. Na figura 4.26(b), está representada a projeção gerada pela aplicação Summit Evolution, relativamente ao edifício que se encontra na figura 4.26(a).



(a) Polígono identificado na fotografia



(b) Projeção 3D do polígono

Figura 4.26: Representação do polígono a 2D e 3D

4.3.6 Outros processos

Nesta secção será abordado um conjunto de técnicas que foram testadas, no decorrer desta dissertação, cujos resultados obtidos não foram os desejados. Alguns devido a não serem direccionados para este objetivo e outras devido à sua complexidade.

4.3.6.1 Interseção direta

A interseção direta é um processo topográfico, que através de dois pontos conhecidos calcula a localização de um terceiro ponto.

Na fotografia de origem são selecionados dois pontos e medidos os ângulos para o ponto P , que pretendemos determinar na fotografia de destino. De seguida é calculado o azimute para o ponto P , conforme se encontra na figura 4.3.

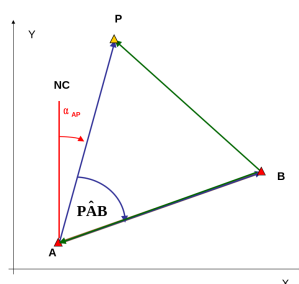


Figura 4.27: Interseção direta

Através da equação 4.3 é possível determinar a localização do ponto P .

$$\begin{aligned} X_P &= X_A + \overline{AP} \sin \alpha_{AP} \\ Y_P &= Y_A + \overline{AP} \cos \alpha_{AP} \end{aligned} \quad (4.3)$$

Este método foi implementado, mas possui dois problemas que o inviabilizam. O primeiro prende-se com o facto de não ser possível determinar um referencial comum

para medir o azimuth, pois as fotografias possuem translações diferentes. O segundo é devido à não preservação dos ângulos entre fotografias.

Este método acabou por ser abandonado.

4.3.6.2 Ceviana

Uma Ceviana é um segmento de reta, num triângulo, que passa num vértice e intersecta o lado oposto. O teorema de Ceva²⁰ implica que a interseção das três cevianas concorrentes, determinam um único ponto, denominado ponto Ceviano[EWa], podemos visualizar uma representação destes segmentos de reta e os respectivos pontos no triângulo na figura 4.28.

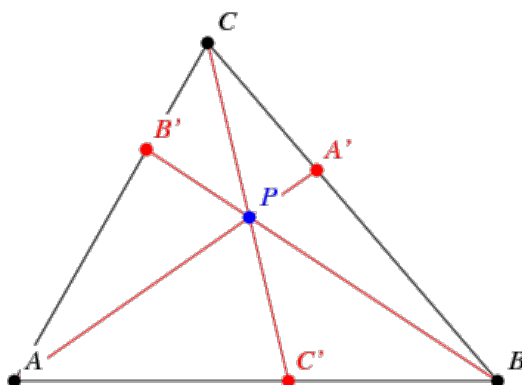


Figura 4.28: Cevianas

A aproximação implementada para detetar os pontos homólogos, necessita de seleccionar o triângulo onde o ponto está inserido, na fotografia de origem. De seguida passa-se as cevianas por esse ponto. Ao fazermos a interseção entre o segmento de reta da ceviana e o lado oposto, determinamos os pontos apresentados na figura 4.28, A' , B' , C' . Com estes pontos são calculadas as distâncias dos dois segmentos de reta resultantes, em cada lado do triângulo, eg: \overline{AC} e $\overline{B'C}$.

Na fotografia de destino, selecciona-se o mesmo triângulo e calcula-se a distância do segmento \overline{AC} , no triângulo de destino.

Fazendo a relação entre as duas distâncias do segmento \overline{AC} , nos dois triângulos, e a distância $\overline{B'C}$ no triângulo de origem, determina-se o valor de $\overline{B'C}$, no triângulo de destino e a respetiva localização do ponto B' . Este processo é repetido para cada um dos lados do triângulo.

Basicamente estamos a reconstruir as cevianas, traçadas no triângulo de origem, no triângulo de destino, fazendo a interseção das três cevianas, determinamos o ponto homólogo. Se os segmentos não se intersectarem no mesmo ponto, é determinado um triângulo e o seu centro é considerado o ponto homólogo.

Este método utiliza somente distâncias e a relação entre elas, não fazendo recurso ao cálculo de ângulos, devido a estes não se manterem entre fotografias.

²⁰Provado em 1678 por Giovanni Ceva

Apesar de ter sido feita a implementação, o método apresentou erros na localização do ponto homólogo, sobretudo no interior de triângulos onde o terreno era mais acidentado e acabou por ser abandonado antes da fase de testes.

4.3.6.3 Retificação/normalização das fotografias

A retificação/normalização das fotografias, de um modelo estereoscópico, visa uma série de transformações, em ambas as fotografias, de forma a que os pontos homólogos se localizem sobre a mesma paralela a um dos eixos.

Quando as fotografias aéreas são tiradas, a câmara não se mantém sobre o mesmo eixo, as transformações inerentes à retificação/normalização permitem que sejam calculado um novo par de fotografias, alinhadas sobre o mesmo eixo.

A figura 4.29, apresenta em forma de diagrama, as tarefas necessárias para realizar a retificação das fotografias do modelo estereoscópico. A verde encontram-se as funções que o OpenCV disponibiliza, com os respetivos dados de entrada e saída. A azul as matrizes e a laranja os vectores que são necessários determinar. A Branco são representados os ficheiros.

Para efetuar este processo é inicialmente necessário efetuar a calibração das duas câmaras. O OpenCV possui uma função para fazer este cálculo, que necessita como dados de entrada:

- Pontos homólogos²¹ origem: vector com a localização dos pontos homólogos na fotografia de origem;
- Pontos homólogos²¹ destino: vector com a localização dos pontos homólogos na fotografia de destino;
- Coordenadas terreno: vector com as coordenadas terreno dos pontos homólogos das fotografias;
- Matriz da câmara origem: matriz com os parâmetros internos da câmara utilizada para tirar a fotografia origem;
- Matriz da câmara destino: matriz com os parâmetros internos da câmara utilizada para tirar a fotografia destino²²;
- Distorção da câmara origem: vector com os valores de distorção da câmara origem²³;
- Distorção da câmara destino: vector com os valores de distorção da câmara destino^{22,23};

²¹Identificados no ISAT.

²²Neste caso a câmara é a mesma e estes valores são os mesmos da câmara origem.

²³Os vectores de distorção podem ser ajustados na calibração, e utilizados a partir deste momento.

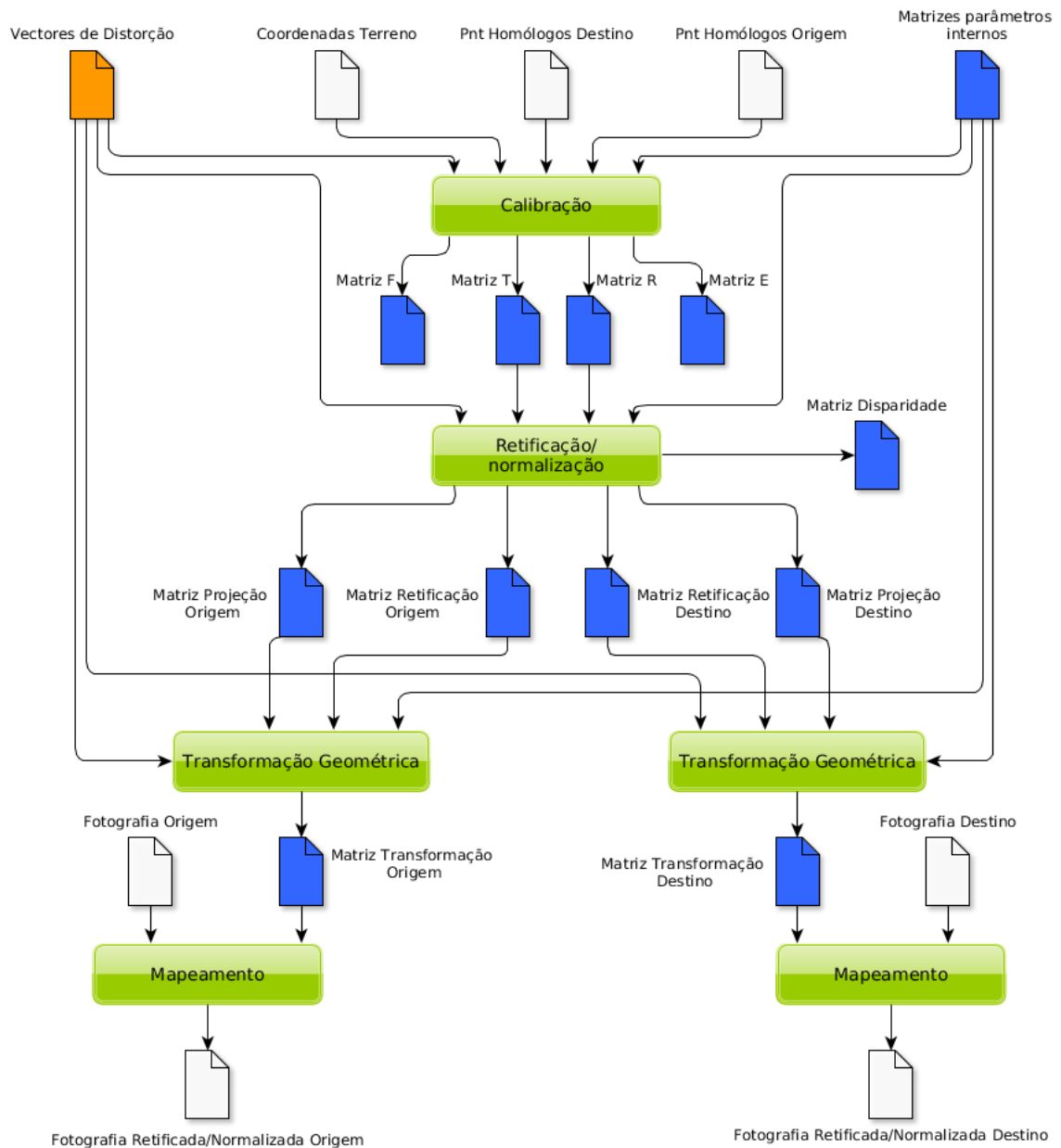


Figura 4.29: Diagrama da retificação/normatização fotográfica

Esta função está adaptada para ser utilizada na captação de fotografias, com duas câmaras, lado a lado, normalmente utilizadas em projetos de investigação e recentemente introduzidas na indústria automóvel.

Normalmente esta calibração é feita com o recurso a um padrão xadrez, onde são identificados os pontos homólogos, entre as duas fotografias, e as coordenadas terreno, encontram-se num plano, perfeitamente definidas²⁴.

No caso das fotografias utilizadas nesta dissertação, as câmaras já se encontram calibradas e as distorções eliminadas. A calibração é executada para ter acesso às matrizes produzidas por este processo, permitindo assim utilizar a função de retificação/normalização disponível no OpenCV.

No caso concreto deste trabalho houve a necessidade de recorrer aos pontos provenientes da aerotriangulação, para termos acesso às coordenadas terreno. Esta restrição faz com que todo o processo só possa ser executado nas fotografias completas. Devido a estas terem um elevado tamanho, este processo é demorado.

O processo de calibração retorna como dados de saída um conjunto de matrizes:

- R: matriz de rotação entre o sistema de coordenadas entre a fotografia origem e fotografia destino;
- T: vector de translação entre o sistema de coordenadas das câmaras;
- E: matriz essencial;
- F: matriz fundamental.

A retificação/normalização utiliza as matrizes T e R, da calibração, e as matrizes das câmaras e distorção, como dados de entrada, para produzir uma matriz de retificação/-normalização e projeção para cada fotografia.

Para além dos pares de matrizes, esta função produz uma matriz para efetuar o mapeamento de disparidades com a profundidade.

Com o respetivo conjunto de matrizes, e as matrizes das câmaras e distorção, são calculadas as transformações geométricas de cada fotografia. Esta função vai calcular as transformações que são necessárias, para eliminar a distorção e retificar/normalizar a fotografia.

De seguida é necessário efetuar um mapa de transformação para cada fotografia e o respetivo mapeamento dos píxeis das fotografias originais, nas novas fotografias retificadas/normalizadas e sem distorção.

A retificação/normalização das fotografias facilita a pesquisa de pontos homólogos, pois neste momento estes encontram-se alinhados com um dos eixos, no entanto surge o problema da nova localização dos píxeis. Para o solucionar é necessário recorrer a transformações entre a localização do píxel na fotografia origem para a fotografia origem retificada, efetuar a pesquisa na fotografia destino retificada, e efetuar a transformação inversa, para obter o píxel homólogo na fotografia de destino.

²⁴Pois é conhecido o tamanho dos quadrados do xadrez.



Avaliação

Neste capítulo efetuamos a avaliação dos resultados obtidos com a implementação do método apresentado nesta dissertação, que foi abordado no capítulo 4.

Inicialmente é efetuada uma descrição do tipo de zonas selecionadas, onde é aplicado o método que propomos, descrevendo as características principais de cada uma delas. Estas visam representar os as zonas mais comuns encontradas no território nacional.

Em termos de resultados quantitativos, pretendemos avaliar a identificação de edifícios, sobre dois critérios, com o primeiro pretendemos avaliar o número de edifícios identificados. No segundo avaliamos as áreas identificadas dos edifícios. Em seguida avaliamos a extração em termos da qualidade apresentada pelos resultados.

A avaliação qualitativa pretende avaliar os resultados obtidos em termos de aceitação, segundos as normas estabelecidas, tendo em atenção a necessidade de efetuar uma avaliação em termos da qualidade dos polígonos gerados em 2D e 3D.

Os resultados são avaliados em relação a um conjunto de edifícios previamente restituídos por um operador, na aplicação *Summit Evolution*, e que consideramos como sendo os corretos para a elaboração da avaliação do método que apresentamos.

5.1 Áreas avaliadas

As fotografia aéreas que foram fornecidas para a elaboração desta tese, são propriedade da DGT tendo sido obtidas em 2012 na zona de Braga. Estas fotografias foram tiradas com a câmara UltraCam Eagle, cujas características podem ser consultadas na tabela 2.1. Cada fotografia tem uma dimensão de 13080×20010 píxeis, representando aproximadamente uma área de $6800 \times 10400 \text{ m}^2$ e ocupando 750 Mb cada. Para a implementação do protótipo, são utilizados segmentos destas fotografias, que se caracterizam por conterem

edifícios diferenciados entre zonas. Desta forma retiramos de cada fotografia original 1% de cada tipo de zona. Assim ficamos com uma fotografia com 1308×2001 píxeis, analisando uma área aproximada de $680 \times 1040 \text{ m}^2$.

Cada fotografia avaliada, é identificada pela palavra *set* e uma letra, seguida de dois algarismos para identificar o tipo de zona e três algarismos para identificar a fotografia original de onde foi retirado o segmento¹. Enquanto que a letra e a identificação da zona servem somente para organização, os três algarismos da fotografia original, permitem seleccionar os parâmetros de orientação externa que serão utilizados, de forma a aplicar nas equações de colinearidade. Assim a fotografia *seta01534.tif* é o primeiro extrato que representa a zona residencial, esta foi retirada da fotografia original que termina em 534. No caso de haver a necessidade de seleccionar outra zona residencial da fotografia 534 a letra é incrementada ficando a fotografia com a designação *setb01534.tif*.

Foram seleccionados quatro zonas que representam, zonas residenciais(01), zonas urbanas(02), zonas históricas(03) e zonas rurais(04). Para estas foram restituídos todos os edifícios por um operador de fotogrametria, sendo estes utilizados para efetuar comparações com os resultados deste método. As áreas restituídas são identificadas pela abreviatura *set*, pela letra de identificação e os algarismos da identificação da zona. Na tabela 5.1 apresenta-se o número e a área em m^2 dos edifícios de cada zona. Este valores serão representados nos gráficos por uma linha horizontal a preto.

Tabela 5.1: Número de píxeis em cada máscara

	seta01	seta02	seta03	seta04
Número de edifícios	210	101	69	72
Área dos edifícios	38776	96425	191824	11695

Através da tabela 5.1, podemos verificar que a fotografia *seta03*, que corresponde à zona histórica, apesar de ser aquela que apresenta o menor número de edifícios restituídos, possui a maior área restituída. Este facto deve-se às normas de aquisição [IGe08], pois determinam que edifícios que se encontrem juntos são restituídos no mesmo polígono. Estas características também se encontram presentes na fotografia do tipo de zona urbana *seta02*.

5.1.1 Zona residencial

Nesta zona pretendemos identificar edifícios caracterizados por possuírem um ou dois pisos, estando estes separados uns dos outros e com contraste entre a cobertura e o terreno.

A fotografia utilizada desta zona é apresentada na figura 5.1(a).

¹Os três algarismos são únicos para cada fotografia



(a) Ficheiro seta01534.tif



(b) Ficheiro seta02534.tif



(c) Ficheiro seta03536.tif



(d) Ficheiro seta04585.tif

Figura 5.1: Imagens das zonas avaliadas

5.1.2 Zona urbana

Nesta zona pretendemos identificar edifícios caracterizados por possuírem mais que dois pisos, constituindo construções em altura, separados ou juntos uns dos outros e com baixo contraste entre a cobertura e o terreno.

A fotografia utilizada desta zona é apresentada na figura 5.1(b).

5.1.3 Zona histórica

Nesta zona pretendemos identificar edifícios caracterizados por serem baixos, com ligação entre eles, feita de forma irregular e apresentando bom contraste entre edifícios e o terreno. Com uma separação entre os enfiamentos dos edifícios reduzida.

A fotografia utilizada desta zona é apresentada na figura 5.1(c).

5.1.4 Zona rural

Nesta zona pretendemos identificar edifícios caracterizados por serem baixos, sem ligação e com contraste entre edifícios e o terreno.

A fotografia utilizada desta zona é apresentada na figura 5.1(d).

5.2 Taxa de identificação

Nesta secção pretendemos analisar a taxa de identificação, que o método possui na identificação de edifícios, apresentando para este efeito duas abordagens diferenciadas. Na primeira iremos apresentar o número de edifícios que é possível identificar, não tendo em atenção a forma ou a quantidade do edifício que é identificado. Na segunda pretendemos identificar a área que é possível identificar, neste caso já será tomado em consideração a quantidade do edifício foi identificada.

De forma a verificar qual o melhor espaço de cores, para utilizar neste método, a análise da taxa de identificação é realizada sobre os seis espaços de cores referidos na secção 4.1.3.1. No entanto para o espaço HSV utilizamos duas máscaras, a primeira com o mínimo de cores onde os edifícios estejam separados dos restantes objetos e a segunda com o máximo de cores que é possível utilizar mas os edifícios não devem ficar em dois *clusters* diferentes, ou seja telhados com duas cores.

Recorremos desta forma à aplicação *Quantum Geographic Information System* (QGIS) para efetuar a avaliação destes valores. Esta é uma ferramenta *open source* que permite a visualização, edição e análise de informação geoespacial.

Os polígonos que representam edifícios são classificados em 3 classes, positivos, falsos positivos e falsos negativos.

São considerados como positivos todos os polígonos gerados, que intersetem os reais. Para obter esta informação foi utilizada o comando QGIS apresentado na equação 5.1, para determinar este valor.

$$Positivos = Clip(\text{polígonos gerados}, \text{polígonos reais}) \quad (5.1)$$

São considerados como falsos positivos todos os polígonos que façam parte do conjunto de polígonos gerados mas não intersetem nenhum polígono do conjunto dos reais. Foi utilizada a equação 5.2 para determinar o valor dos falsos positivos. Estes são os polígonos que são gerados pelo método que não existem na realidade, podem ser considerados como aqueles onde o método se “enganou”.

$$Falsospositivos = \text{Polígonos gerados} - \text{Positivos} \quad (5.2)$$

Os falsos negativos são todos os polígonos que façam parte do conjunto de polígonos reais que não intersetem nenhum polígono do conjunto dos gerados, estes são calculados com base na equação 5.3, podendo ser considerados como aqueles que o método “não encontrou”

$$Falsosnegativos = \text{Polígonos reais} - \text{Positivos} \quad (5.3)$$

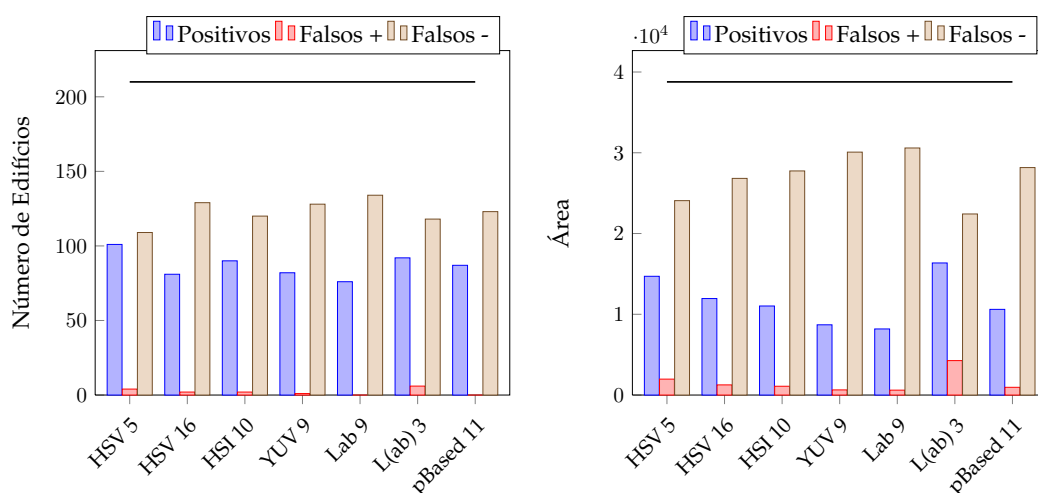
Conforme já foi referido, na análise do número de identificações pretendemos saber quantos polígonos foram gerados pelo método, que correspondem a edifícios reais. Para esta análise não temos em consideração o tamanho ou forma do polígono. No entanto cada polígono só pode identificar um edifício. Se existir um polígono que seja suficientemente grande para agrupar dentro de si dois edifícios, somente é contabilizada uma identificação.

A análise da área identificada tem em consideração se o polígono gerado tem uma maior ou menor capacidade de recolher toda a informação do edifício.

Efetuamos a análise de cada zona em separado, colocando nas figuras apresentadas na secção seguinte, do lado esquerdo encontra-se a classificação da identificação do número de edifícios, enquanto do lado direito está a classificação da área identificada em m².

5.2.1 Zona residencial

Analisando os resultados apresentados na figura 5.2 podemos verificar que a alteração do espaço de cores não produz alterações substanciais em relação ao número de edifícios, mas em termos de área, diferentes espaços de cores produzem alterações.



(a) Número de edifícios identificados na fotografia- (b) Área dos edifícios identificados na fotografia
fia seta01534.tif seta01534.tif

Figura 5.2: Taxa de identificação na fotografia seta01534.tif

Em ambos os casos, a predominância é para os falsos negativos, conforme se apresenta na figura 5.3, a segmentação somente identifica os edifícios que se encontram salientadas a vermelho, na figura 5.3(b). Todos os restantes edifícios são considerados falsos negativos.

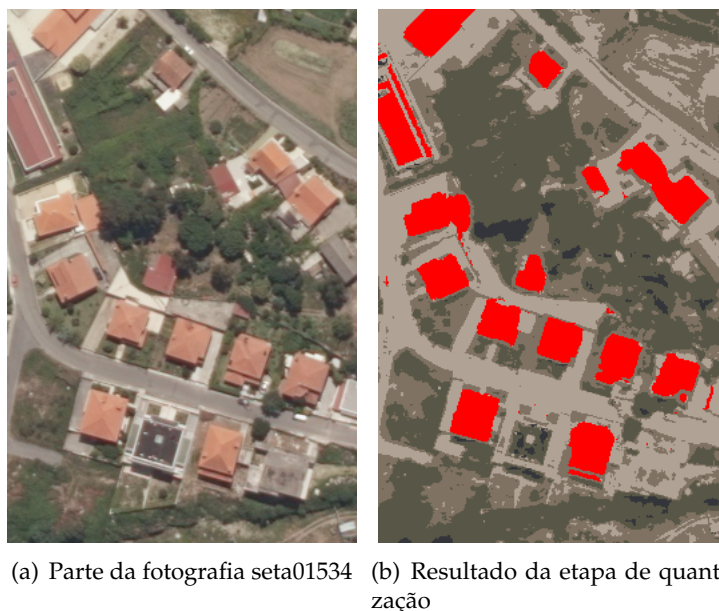


Figura 5.3: Extrato e respetiva quantização da fotografia seta01534

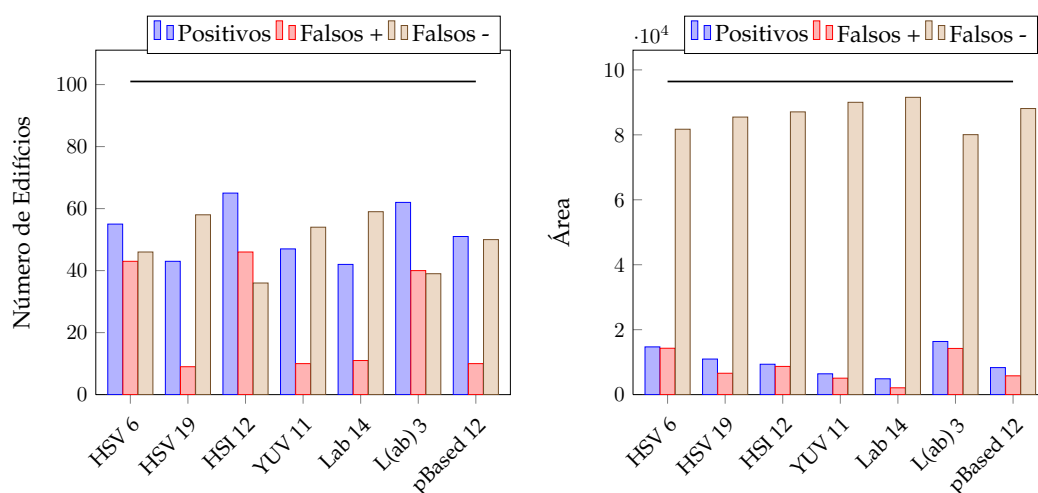
Em termos de polígonos classificados como falsos positivos, o seu número e área acaba por ser reduzida em relação aos que são corretamente identificados. No entanto é importante salientar o reduzido número de falsos positivos apresentado pelo Lab 9 que não se verifica em termos de área. Apesar do Lab 9 ter zero falsos positivos em termos de

número, a área dos falsos positivos não é zero. Isto acontece porque usando este espaço de cores, o método constrói polígonos maiores que as coberturas dos edifícios, sendo estas áreas contabilizadas para falsos positivos. Este pode ser um indício para que sejam utilizados espaços de cores para detetar o edifício e outros para delimitar a área do edifício.

Em termos gerais consideramos que o espaço de cores que melhor efetua a identificação dos edifícios na zona residencial é o HSV com o mínimo de cores possíveis para efetuar a segmentação dos edifícios. Esta seleção surge da combinação dos valores do maior número de positivos com o menor número de falsos positivos.

5.2.2 Zona urbana

Analisando os resultados apresentados podemos verificar na figura 5.4, que a alteração do espaço de cores não produz alterações substanciais em relação ao número e área dos edifícios.



(a) Número de edifícios identificados na fotografia seta02534.tif (b) Área dos edifícios identificados na fotografia seta02534.tif

Figura 5.4: Taxa de identificação na fotografia seta02534.tif

Em termos do número de identificações positivas, podemos verificar que existe um valor significativo de polígonos que identificam edifícios. No entanto quando passamos a analisar as áreas, estes valores indicam que as áreas positivas encontra-se muito próximas das áreas onde o método se "enganou". Em termos de restituição, os edifícios que se encontram juntos são restituídos no mesmo polígono. Nesta zona acontece que alguns edifícios são identificados por vários polígonos.

Na figura 5.5(b), encontra-se salientado a vermelho os edifícios selecionados, na fase de segmentação. Em termos de resultados podemos ver na figura 5.5(c), em azul os edifícios que foram restituídos e a vermelho os polígonos que foram gerados. Aqui podemos verificar que os edifícios que foram restituídos no mesmo polígono são identificados por um

ou mais dos polígonos gerados.

Na zona urbana as coberturas dos edifícios passam a possuir cores mais diversificadas. Na figura 5.5(a), podemos verificar a existência de coberturas cor de laranja e cinzentas. No processo de quantização. Os edifícios com coberturas cinzentas, ficam no mesmo *cluster* das vias, devido à semelhança da sua cor. Se o operador selecionar, para a elaboração da máscara, esta cor, as vias são classificadas como sendo edifícios criando falsos positivos.

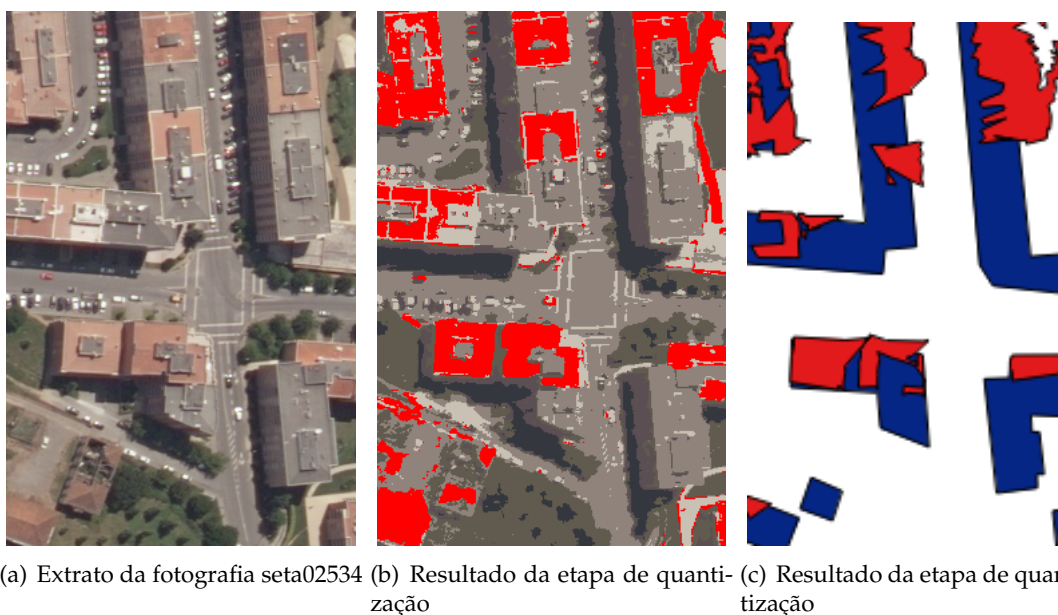
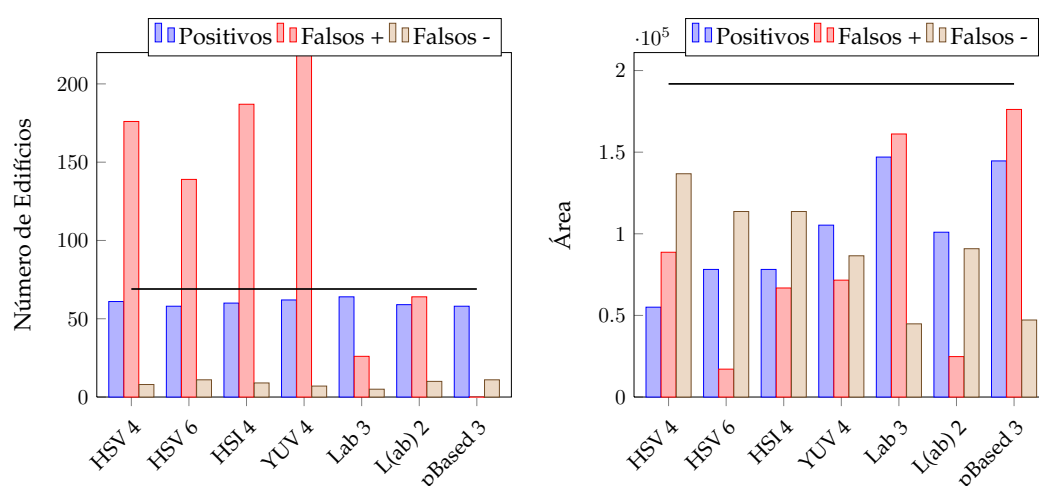


Figura 5.5: Extrato e respetiva quantização da fotografia seta02534

Com a incapacidade de efetuar a separação dos objetos, a aplicação deste método não permite a extração da informação disponibilizada nas fotografias fornecidas. É neste fator que reside a grande diferença entre o número de edifícios não encontrados e a área.

5.2.3 Zona histórica

Ao analisar a figura 5.6, podemos verificar que os espaços de cores apresentam diferenças substanciais entre si, em termos de número e área. Na zona selecionada, devido às coberturas terem cores muito semelhantes e estas encontrarem-se juntas umas das outras, são criados polígonos que englobam vários edifícios, fazendo com que estes sejam de grande dimensão. Nos espaços destes polígonos são ainda criados um elevado número de pequenos polígonos, que grande parte das vezes identificam novamente o mesmo edifício.



(a) Número de edifícios identificados na fotografia seta03536.tif (b) Área dos edifícios identificados na fotografia seta03536.tif

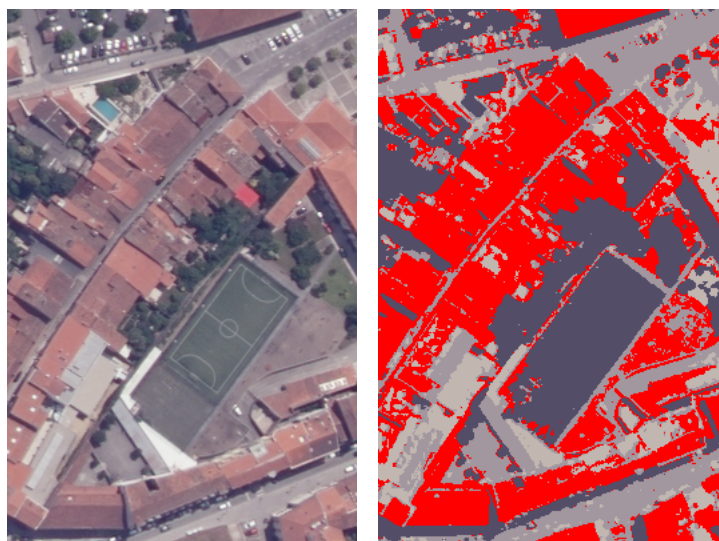
Figura 5.6: Taxa de identificação na fotografia seta03536.tif

Se analisarmos a classificação de falsos positivos, no espaço de cores Lab 3 e PBased 3, verificamos que estes são relativamente baixos em relação aos outros. No entanto, estes espaços de cor em termos de área identificada apresentam os valores mais elevados.

Os polígonos são gerados a partir da segmentação da fotografia, se esta não tiver capacidade para separar os edifícios dos restantes objetos, o método vai apresentar erros. Na figura 5.7(b), podemos verificar que existe uma rua que atravessa a figura na diagonal, esta devido a ser estreita, faz com que existam píxeis que unam as duas áreas. Nesta podemos também verificar que as pequenas áreas que se encontram do lado direito e em cima, na mesma figura, criam pequenos polígonos que identificam, em termos de número, este edifício várias vezes.

O efeito encontrado na zona histórica é inverso ao da zona urbana. Aqui temos presente um elevado número de coberturas de edifícios cor de laranja, que fazem com que existam vários *clusters* diferentes com tonalidades semelhantes. Cores menos representativas com uma piscina, uma viatura ou uma área do solo, ficam no mesmo *cluster* das coberturas dos edifícios, se estas não forem excluídas pelo seu tamanho geram falsos positivos.

Com a análise dos resultados da figura 5.6, que apresentam valores elevados no número de identificações e na área identificada e a apresentação da figura 5.7, que possui um conjunto de áreas muito irregulares, consideramos que este método não possibilita a aquisição de informação com qualidade no conjunto de fotografias disponibilizadas.

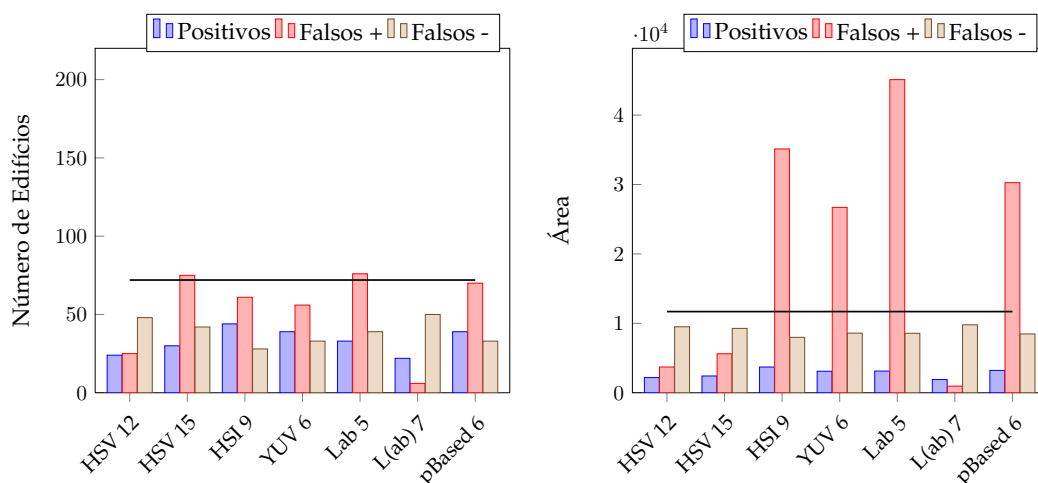


(a) Extrato da fotografia seta03536 (b) Resultado da etapa de quantização

Figura 5.7: Extrato e respetiva quantização da fotografia seta03536

5.2.4 Zona rural

Em termos de espaços de cores, aquele que melhores resultados proporciona é o L(ab) 7, os restantes apresentam valores de falsos positivos elevados em termos de número e área identificada. Como este tipo de zona possui um baixo número de edifícios e de pequenas dimensões, faz com que a cor que representa os telhados seja pouco significativa.



(a) Número de edifícios identificados na fotografia seta04585.tif (b) Área dos edifícios identificados na fotografia seta04585.tif

Figura 5.8: Taxa de identificação na fotografia seta04585.tif

A área dos polígonos classificados como positivos, têm uma pouca expressão em

qualquer dos espaços de cor utilizados, sendo o valor de falsos positivos, em comparação, elevada. O L(ab) 7, destaca-se por ter um valor de positivos superior aos falsos positivos. No entanto, este espaço de cor, para cada conjunto de polígonos positivos gera 50% de falsos positivos, em termos de área identificada.

Devido à pouca representatividade dos edifícios nesta zona, a cor que identifica as coberturas, acaba por ser absorvida por outros *clusters*. Na figura 5.9(b), podemos verificar que as coberturas dos edifícios se encontram no mesmo *cluster* dos terrenos envolventes, este *cluster* encontra-se salientado a vermelho.



(a) Extrato da fotografia seta04585 (b) Resultado da etapa de quantização

Figura 5.9: Extrato e respetiva quantização da fotografia seta04585

O facto de existirem poucos píxeis, que representem a cobertura dos edifícios, nesta zona poderá ser resolvido "forçando" a cor de um ou vários *clusters*. É necessário contudo selecionar a cor para "forçar" o *cluster*. Na elaboração deste método não seguimos esta abordagem devida à dificuldade de encontrar a cor ou intervalo de cores que representam as coberturas dos edifícios. Desta forma consideramos que o método apresentado não se encontra adaptado para efetuar a aquisição nestas zonas.

5.3 Análise pelas fases do método

Iremos nesta secção efetuar a análise do resultados apresentados pelo método, para um espaço de cor. Esta tem por objetivo identificar qual das fases ou etapas deve ser melhorada para obter um incremento no número de polígonos, classificados como positivos.

Conforme foi referido na secção anterior, a zona para a qual o método melhor se adapta é a residencial. Assim selecionamos a quantização realizada com o espaço de cor HSV e cinco cores para efetuar esta análise. O resultado desta quantização encontra-se no anexo C na figura C.1, para melhor ser visualizado. Neste anexo é também incluída a

máscara resultante da fase de segmentação, na figura C.2.

Na fase de segmentação da fotografia são geradas 558 áreas identificadas com a cor selecionada pelo operador.

Manualmente foi feita a classificação das áreas produzidas pela segmentação, que se encontram apresentadas na tabela 5.2. Esta teve em atenção que um edifício só pode ser identificado por uma área da segmentação, existindo assim 71 áreas se não são contabilizadas na tabela.

Tabela 5.2: Classificação das áreas da segmentação

	Positivos	Falsos Positivos	Falsos Negativos
Áreas da segmentação	141	346	69

A fase de construção de polígonos inicia-se com a geração de contornos e eliminação, daqueles que possuem uma área mínima que não pretendemos processar.

Fazendo uma análise aos contornos que são eliminados por área, podemos verificar na figura 5.10, que a área pela qual o maior número de contornos é eliminada, situa-se nos valores mais baixos. Tendo a sua maior expressão na eliminação de área com um píxel.

Para a fase de determinação de arestas e vértices passam um total de 190 contornos, sendo que 24 são contornos que identificam duas vezes o mesmo edifício e não se encontram na tabela 5.3.

Tabela 5.3: Classificação dos contornos

	Positivos	Falsos Positivos	Falsos Negativos
Contornos eliminados	1	320	
Contornos não eliminados	140	26	
Total	140	26	70

Com a figura 5.10, pretendemos demonstrar que esta etapa, tem um papel mais relevante na eliminação de ruído, do que propriamente a eliminação de edifícios de pequenas dimensões. No entanto existe um edifício que é eliminado. Na figura 5.11, apresentamos este edifício e o seu respetivo contorno. Este edifício no entanto apresenta um contorno demasiado pequeno para que seja possível gerar um polígono que o identifique corretamente.

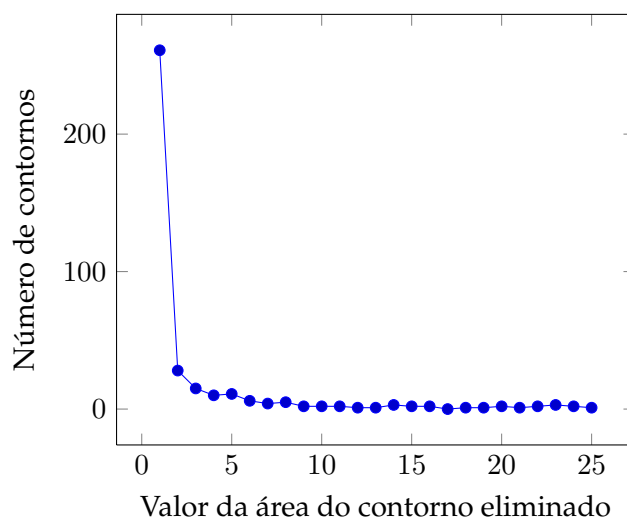


Figura 5.10: Contornos eliminados



Figura 5.11: Contorno positivo eliminado

Na figura C.3 que se encontra no anexo C, podem ser consultados todos os contornos, gerados na construção de edifícios. A vermelho encontram-se os contornos que são eliminados e a azul os restantes.

As etapas relativas à deteção de arestas e vértices produzem os polígonos a 2D, na fotografia de origem. Durante estas etapas, se não for possível extrair três vértices da informação do contorno, este não dá origem a um polígono, passando a ser contabilizado como falso negativo. Na figura 5.12, apresentamos dois exemplos, que representam contornos de áreas positivas, mas que não iram dar origem a polígonos positivos, devido a não se inferir a posição de uma das aresta em falta. Nesta figura, podemos visualizar do lado esquerdo o contorno, ao centro as arestas detetadas e do lado direito os vértices.

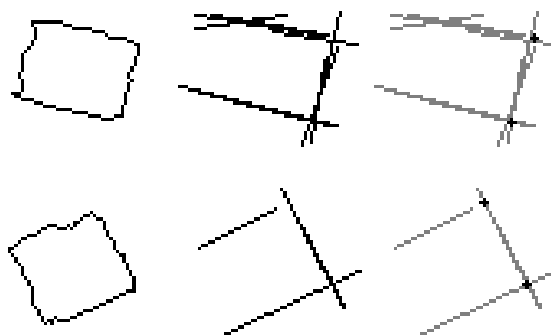


Figura 5.12: Exemplos de polígonos falsos negativos

Na tabela 5.4, apresentamos a classificação dos polígonos no final da fase de construção de polígonos. De salientar a redução do número de falsos positivos, que não foram excluídos na etapa de eliminação de contornos. Apesar de possuírem uma área superior ao mínimo definido, não têm informação suficiente para que seja possível a detecção de arestas e vértices, acabando por ser excluídos.

Tabela 5.4: Classificação dos polígonos 2D

	Positivos	Falsos Positivos	Falsos Negativos
Polígonos	114	4	96

Incluímos na figura 5.13, a representação dos quatro falsos positivos que se encontram mencionados na tabela 5.4, estando ao centro o contorno da área e à direita o polígono resultante da detecção de arestas e vértices.

Na figura C.4 do anexo C, encontram-se desenhados a vermelho os polígonos que resultam da fase de construção de polígonos.

Na fase da determinação de coordenadas terreno, existe a possibilidade de haver alterações na classificação dos polígonos. Devido aos vértices homólogos, serem inicialmente determinados, através de uma transformação afim, que utiliza a triangulação de Delaunay. Se os vértices não se encontrarem dentro do interior de um triângulo, estes não são determinados. Assim, todos os polígonos 2D para os quais não seja possível determinar três vértices 3D, são excluídos. Apresenta-se a classificação destes polígonos na tabela 5.5, onde são perdidos 13 polígonos positivos.

Tabela 5.5: Classificação dos polígonos 3D

	Positivos	Falsos Positivos	Falsos Negativos
Polígonos	101	4	109

Na figura C.5 do anexo C, a vermelho representamos os polígonos que se encontram fora da área onde existe triangulação, esta a azul os restantes. Nesta figura as linhas a



Figura 5.13: Exemplos de polígonos falsos positivos

branco representam a aplicação do algoritmo de Delaunay.

Para finalizar esta secção apresentamos a tabela 5.6, um resumo das tabelas anteriores. Em azul estão representadas as diferenças em termos da classificação introduzidas por cada uma das fases.

Tabela 5.6: Classificação dos polígonos

	Positivos	Falsos Positivos	Falsos Negativos
<i>Input</i> da segmentação da foto	0	0	210
Alterações	+141	+71	-69
<i>Output</i> da segmentação da foto	141	71	69
<i>Input</i> da construção de polígonos	141	71	69
Alterações	-27	-67	+27
<i>Output</i> da construção de polígonos	114	4	96
<i>Input</i> da determinação de coordenadas terreno	114	4	96
Alterações	-13	0	+13
<i>Output</i> da determinação de coordenadas terreno	101	4	109

Inicialmente o número de falsos positivos é aproximadamente 50% do número de positivos, sendo este valor reduzido na fase da construção de polígonos. Devido à eliminação de contornos com pequena área e à dificuldade de encontrar arestas nestes contornos.

O número de positivos encontrados na fase da segmentação da fotografia é de aproximadamente 67%, do número total de edifícios. Devido às coberturas dos restantes edifícios não apresentarem as mesmas características radiométricas, não foi possível efetuar a sua identificação.

A fase responsável por "perder" mais positivos é a da construção de polígonos, esta não efetuar a análise das características recolhidas, sendo assim obrigada a excluir polígonos.

Devido ao método estar limitado à área que possui triangulação, existem polígonos para os quais não será possível determinar a sua altimetria.

5.4 Taxa de aceitação

A taxa de aceitação tem por objetivo analisar a qualidade dos polígonos criados pelo método apresentado nesta dissertação. Para aferir os resultados recorreremos ao STANAG [22102], que classifica a informação geográfica para a produção de cartografia. Esta classificação é realizada, por níveis, estando a informação planimétrica e altimétrica separadas em termos de classificação, encontrando-se a informação restituída pelo IGeoE no nível mais elevado.

Para efetuar a comparação dos resultados recorreremos à aplicação ArcGIS, da empresa *Environmental Systems Research Institute*, esta possui licenciamento, tendo sido permitida

a sua utilização para a elaboração desta tese.

Foi efetuada a comparação dos edifícios restituídos com os gerados, utilizando o espaço de cores HSV com 5 cores, visto ter sido esta a configuração que apresentou os melhores resultados, em termo de taxa de identificação.

Como nesta análise pretendemos saber quantos seriam os polígonos aceites para a elaboração de cartografia, a forma é um requisito necessário. Assim, utilizamos para esta comparação os polígonos gerados que se assemelham com os restituídos, tendo sido eliminados os restantes. Ficamos com 33 polígonos em ficheiros separados. Nas figuras D.1 e D.2, podem ser visualizados os polígonos e vértices de cada um dos ficheiros.

Utilizando o ArcGIS, extraímos de cada polígono os seus vértices e as respetivas coordenadas 3D. Em seguida utilizamos o comando *spatial join*, disponibilizado na *toolbox* da aplicação. Este agrupa cada vértice restituído com o vértice mais próximo gerado e disponibiliza uma tabela com essa informação.

Esta tabela foi exportada para uma folha de cálculo. Nesta foram calculadas as distâncias dos vértices agrupados, pelos atributos (x, y) e z , sendo de seguida calculada a média destes valores. A média correspondente à planimetria, foi de 1.7 m. Segundo o STANAG [22102], esta fica classificada no nível A, o mais elevado. No que diz respeito à altimetria, o valor da média apurado foi de 3.95 m, fazendo com que esta informação fique classificada no nível 1, o segundo da classificação. A frequência das distâncias entre vértices, em planimetria e altimetria, encontram-se apresentadas respetivamente nas figuras 5.14(a) e 5.14(b), a linha vermelha representa a tendência destes valores.

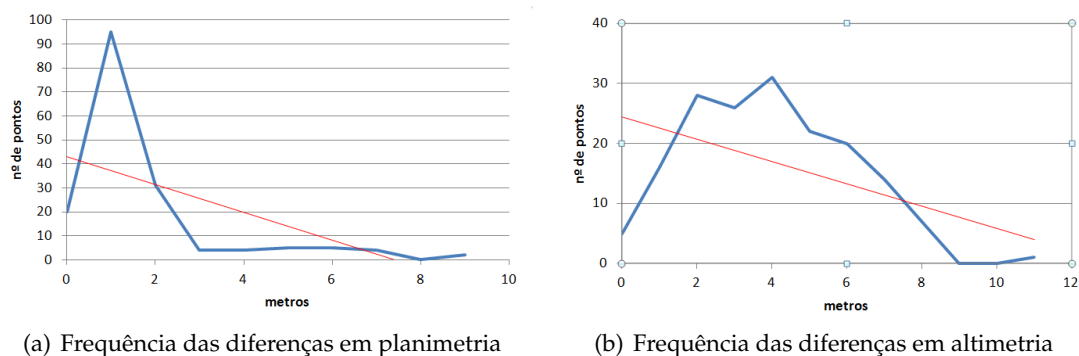


Figura 5.14: Frequência das diferenças

Conforme se pode verificar no resultado da média e na frequência dos valores de distância obtidos, em termos planimétricos os valores estão completamente dentro das normas², impostas pelo STANAG [22102].

Em termos altimétricos, estes valores não se encontram no nível de classificação mais elevado. Quando utilizamos as equações de colineariedade para a determinação das coordenadas 3D, usamos uma coordenada origem, retirada de uma representação vetorial,

²O limite inferior que os valores devem possuir para não passar para o nível de baixo, é de 12.5 m.

e uma coordenada destino determinada numa representação raster. Enquanto que a primeira é um número decimal a segunda é um número inteiro, Se na coordenada destino utilizarmos o número $x.1$ e o $x.9$, no cálculo das equações de colinearidade, o resultado é bastante diferente. Sendo que a maior variação acontece no valor do z . No anexo B, apresentamos um caso concreto desta variação.

Desta forma realizámos uma experiência, que consistiu em gerar três conjuntos de polígonos. No primeiro com a coordenada destino no centro do píxel, no segundo é utilizada a parte fracionaria da coordenada de origem e o terceiro a coordenada destino não utiliza parte fracionaria representando o canto superior esquerdo do píxel, "início" do píxel. Os valor médios obtidos pelos três conjuntos são apresentado na tabela 5.7.

Tabela 5.7: Localização da coordenada destino

Média das Distâncias de Z	
Centro do píxel	3.95
Fracionário	3.79
"Início" do píxel	3.06

Como podemos verificar, em termos médios, a não utilização de parte decimal, diminui a diferencia em termos altimétricos, mas não altera a classificação da qualidade da informação.

Recorremos assim a outra experiência. Esta tem por objetivo, efetuar a comparação de vértices, fazendo a alteração da superfície onde estes se inserem, sendo utilizadas três abordagens. Na primeira calculamos a média da cota de cada polígono gerado, com esse valor determinamos a distância a cada vértice do polígono restituído e calculamos a média destas distâncias. Na segunda colocamos todos os vértices dos polígonos gerados e restituídos, na cota do vértice mais elevado, calculamos a distância de cada um e fazemos a média. Na terceira calculamos a média da cota de cada polígono, seja este gerado ou restituído, e procedemos ao cálculo da média novamente.

Tabela 5.8: Localização dos vértices na superfície

	Media de pnt gerados	Maior cota de reais e gerados	Média de gerados e reais
Centro do píxel	3.47	2.78	4.01
Fracionário	3.38	2.76	3.92
"Início" do píxel	2.56	1.91	3.08

Na tabela 5.8, são apresentados os valores retiradas da segunda experiência. Podemos verificar que se colocarmos todos os vértices na cota do vértice com maior cota, a média das distâncias entre os vértices dos polígonos restituídos e gerados apresentam valores que se permitem a classificação mais elevada segundo o STANAG [22102].

No entanto a utilização de 33 edifícios, não representa uma amostra suficientemente

grande para que se possam tirar conclusões. De referir que em termos de implementação só se encontra disponível a abordagem inicial, com o cálculo das equações de colinearidade que utilizam os valores do centro do píxel e sem alterações nas superfícies do polígono.

Conclusões e trabalho futuro

Para finalizar a elaboração deste documento, apresentamos as conclusões relativas ao método proposto. Este visa efetuar a restituição de edifícios para a produção de cartografia. Apresentamos ainda um conjunto de procedimentos que podem ser importantes, para potenciar a obtenção de melhores resultados, em termos de quantidade e qualidade. São ainda abordadas técnicas que não foram utilizadas mas que, na nossa opinião, podem de ser relevantes para aumentar a performance e substituir etapas apresentadas nesta dissertação.

6.1 Conclusões

Nos objetivos apresentados na secção 1.3, propomos-nos a desenvolver um método com a capacidade de detetar e efetuar a extração de edifícios, com a sua informação altimétrica. Consideramos que o objetivo foi atingido, demonstrando que existe a possibilidade de aplicar um método que utiliza a informação do modelo estereoscópico para realizar esta tarefa, havendo ainda necessidade de explorar outras abordagens no que diz respeito à identificação de edifícios, para que o método seja utilizado de forma sistemática.

O método constitui uma primeira abordagem a esta temática, cria os fundamentos para o desenvolvimento de uma ferramenta com a capacidade de diminuir o tempo despendido na restituição de edifícios, com os recursos existentes na cadeia de produção do IGeoE.

Durante o desenvolvimento, foi ainda possível, verificar que existe a possibilidade de estender este método para outros objetos, que são necessários para a produção de cartografia do IGeoE.

Conforme pretendíamos este método não necessita da informação de modelos digitais, nem outra informação que não esteja presente na cadeia de produção onde se insere.

Verificamos que a utilização do algoritmo de detecção de arestas, referenciado no trabalho relacionado, não permitiu a obtenção de resultados satisfatórios, quando da sua aplicação nas fotografias fornecidas.

A utilização de vários modelos de cores demonstrou, que a forma como a informação se encontra organizada, contribui para a obtenção de melhores resultados.

O processo de quantização permite efetuar o agrupamento de píxeis, mediante a informação radiométrica, presente na fotografia. No entanto, conforme verificámos não foi possível aplica-lo, com sucesso, em todos os tipos de zonas.

Foi possível integrar, no método apresentado, procedimentos que não foram criados com o objetivo para o qual os utilizamos.

Foram exploradas um conjunto de técnicas, que apesar de não se encontrarem neste relatório, contribuíram para perceber o problema. Nomeadamente os filtros *Contrast stretching* e *Bilateral*, o algoritmo de detecção de cantos Harris e Shi-Tomasi, os operadores *Sobel* e *Laplace* e os algoritmos *Watershed* e *Floodfill*.

Em termos de análises realizadas, verificamos que a taxa de identificação, no que diz respeito ao número de edifícios encontrados é satisfatória para a utilização do método com a máscara de uma cor. No entanto em termos de área, ainda existe a possibilidade de melhoramentos. Foi ainda verificado que o método terá algumas dificuldade, em extrair informação de determinados tipos de zonas:

- Zonas em que os edifícios se encontrem juntos e que exista variação de cor nas coberturas;
- Zonas onde não exista um equilíbrio, da informação radiométrica.

A taxa de identificação, resultante da aplicação do método, permitiu alcançar o valor apresentado nas normas, em termos de planimetria. No entanto em altimetria, não foi possível classificar esta informação no nível mais elevado.

Na análise pelas várias fases, verificamos que a construção de polígonos, requer uma atenção particular pois é responsável por um decréscimo do número de polígonos positivos.

6.2 Trabalho futuro

O trabalho futuro que propomos para este trabalho encontra-se dividido em aspetos gerais que são transversais a todo o método e aspetos específicos que incidem sobre de cada uma das fases abordadas.

6.2.1 Aspetos gerais

A utilização e alteração de parâmetros de todo o método, encontra-se desacoplada da implementação, este procedimento constitui uma boa prática de programação, sendo a

nossa intenção, que estes parâmetros, devido ao seu elevado número, devam constituir um ficheiro de configuração em que o operador, com permissões, tenha a capacidade de alterar sem que exista a necessidade de efetuar a compilação da aplicação para estes produzam os seus efeitos.

Para aumentar a performance em termos de tempo existe a possibilidade de paralelizar algumas das etapas do método. Na segmentação é possível paralelizar o processo de geração da máscara, após a seleção da cor pretendida, na etapa da escrita da imagem binária. Na deteção de arestas e vértices, os contornos são processados individualmente, permitindo que possam ser paralelizados. Na determinação de coordenadas terreno, a identificação do ponto na fotografia de destino e o cálculo das equações de colinearidade podem ser feitos em paralelo para cada um dos pontos, mas existe a necessidade de preservar a ordenação dos pontos no polígono.

Este método tem a possibilidade de utilizar informação pré-processada, pois a determinação de pontos homólogos, pela técnica SIFT e SURF, a triangulação dos pontos homólogos, a matriz afim de cada triângulo e a matriz fundamental, de cada modelo estereoscópico podem ser pré-processados. Pois esta informação pode ser calculada para cada modelo, independentemente.

6.2.2 Segmentação da imagem

Uma das características presentes no tipo de fotografias que são utilizadas neste método, é a elevada concentração de informação. Seria benéfico efetuar uma simplificação da mesma, efetuando a eliminação de conjuntos de cores que não contribuam para o objetivo que pretendemos alcançar.

Outra forma de selecionar a máscara pode passar por selecionar a região que tenha a cor mais próxima de um conjunto previamente estabelecido como cores frequentes na cobertura de edifícios. Esta não foi utilizada pois no território nacional não existe regularidade definida globalmente em termos das cores existentes de coberturas.

Em termos de determinação de proximidade de cores, o algoritmo K-means, utiliza a distância euclidiana, visto ser esta a forma que está implementada de origem. Os modelos de cores CIE, inicialmente utilizaram esta metodologia, para calcular a proximidade das cores, tendo sido alterada em 2004 pelo standard ISO/CIE 11664-6:2014-02, que se encontra detalhado no artigo[SWD05].

A quantização de canais selecionados de um espaço de cores, foi apenas utilizada com o espaço Lab. Existe a possibilidade que outros espaços possam ter resultados semelhantes ou melhores.

Existem indícios suficientes que levam a ponderar a possibilidade da utilização de mapas de disparidades, para efetuar a seleção das áreas que pretendemos avaliar. Esta técnica é ligeiramente diferente da que propomos neste método, pois utiliza a informação presente nas duas fotografias para produzir a segmentação. Esta apresenta como desvantagem a sua sensibilidade às elevações do terreno.

6.2.3 Construção de polígonos

Conforme foi referido no capítulo 4 a implementação da classe **Contour** permite guardar a informação relativa às várias características, nomeadamente:

- Estrutura para armazenar a informação dos pontos do contorno 6.1(a);
- Estrutura para armazenar as arestas produzidas pelo algoritmo Hough 6.1(b);
- Estrutura para armazenar as arestas fundidas 6.1(c);
- Estrutura para armazenar as arestas estendidas 6.1(d);
- Estrutura para armazenar os vértices (interseção de arestas) 6.1(e);
- Estrutura para armazenar os vértices fundidos 6.1(f).

Em cada uma das alíneas encontra-se a referência para a figura que representa o conteúdo destas estruturas.

Este trabalho não efetua a análise de toda a informação, pois o seu objetivo encontra-se em representar a informação presente nas fotografias e não o de fazer inferências sobre esta. Esta foi armazenada para facilitar uma possível extensão do trabalho, onde seja feita a análise, podendo desta forma inferir a localização de características. A título de exemplo, na sequência de passos da figura 6.1 existe a possibilidade de inferir um segmento de reta para fechar o polígono, com a informação dos dois vértices da figura 6.1(f) e as arestas que se encontram paralelas da figura 6.1(c).

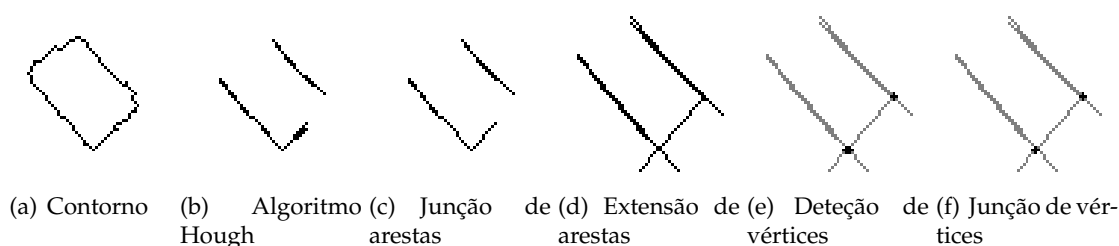


Figura 6.1: Inferir de arestas

O algoritmo apresentado no artigo [GJMR12], apesar de ter sido publicado recentemente, no ano de 2013, apresenta características que podem levar à substituição de várias etapas apresentadas neste método, nomeadamente deteção de contornos, o algoritmo Hough para detetar arestas, a junção de arestas, extensão e a segunda junção. Desta forma seria apenas necessário utilizar as etapas referentes à deteção de vértices.

6.2.4 Determinação de coordenadas terreno

Na correlação de pontos é utilizados o exemplo apresentado no tutorial do OpenCV, para encontrar os pontos, a partir de uma distancia entre vizinhos. Seria relevante utilizar a

abordagem apresentada no artigo [ML09], de forma a verificar quais as diferenças entre ambos.

O coeficiente de correlação utilizados neste método efetua todos os cálculos sobre uma imagem em escala de cinzentos para agregar os canais RGB num valor. Este procedimento pode ser realizado sobre os canais em separado, tendo em atenção que de seguida iremos obter três valores, um para cada canal, sendo necessário efetuar a desambiguação.

Um possível refinamento na determinação das coordenadas terreno é o de recorrer à informação altimétrica, que se encontra em volta do vértice identificado. Mediante os declives existentes ao redor do vértice, existe a capacidade de analisar essa informação para efetuar a recolocação do mesmo.

Existe ainda a possibilidade de estender este trabalho para a identificação de outros objetos que se encontram no catalogo do IGeoE. Um dos objetos que mais se evidenciou durante o trabalho de implementação, foram as piscina, que são catalogados com o símbolo tanque de água. Devido a terem um elevado contraste e radiometria em relação ao que se encontra à sua volta.



Algoritmo probabilístico da transformada de Hough

O algoritmo probabilístico da transformada de Hough, apresentado no artigo [MGK98], permite extrair dos contornos a informação dos segmentos de reta mais significativos.

Esta função disponibilizada pelo OpenCV, utiliza um parâmetro para definir o ângulo pelo qual vai ser feita a "pesquisa" de arestas na imagem. Para esta dissertação foram utilizados valores que variam entre 0° e 90° .

Para percorrer os diversos ângulos, utilizamos um ciclo que faz o incremento do ângulo segundo um intervalo. Tendo sido selecionados os valores 1, 0.1, 0.01 e 0.001 para avaliação.

A avaliação foi realizada com a máscara gerada a partir do ficheiro *seta01534.tif*, sobre o espaço de cor HSV e cinco cores, estando os valores recolhidos apresentados na tabela A.1.

Tabela A.1: Algoritmo da transformada de Hough

	Tempo (seg)	Total de Segmentos	Nº de segmentos não duplicados
Intervalo 1	1.19	12002	2512
Intervalo 0.1	13.69	120830	7438
Intervalo 0.01	166.85	1207638	12502
Intervalo 0.001	2070.84	12072857	16474

A coluna tempo, representa o tempo necessário para executar um ciclo, tendo sido o valor determinado executando cinco vezes o ciclo para cada intervalo. De seguida fazemos a média dos três valores intermédios sendo eliminados o primeiro e o último.

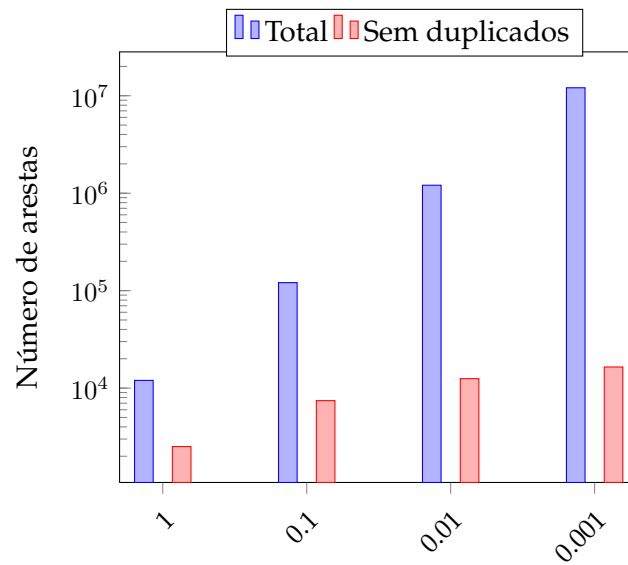


Figura A.1: Relação entre total de segmentos e segmentos não duplicados

Conforme podemos verificar quando o intervalo diminui 10%, o tempo necessário de processamento aumenta mais de 10 vezes. Este valor não se encontra fora das expectativas, pois a diminuição do intervalo faz com que cada ciclo tenha que produzir mais iterações entre os ângulos 0° e 90° .

A coluna total de segmentos representa o número total de segmentos encontrados pelo algoritmo com os diversos intervalos. Na coluna segmentos não duplicados, foram eliminados aqueles que se encontravam duplicados, isto porque de um conjunto de píxeis podem ser extraídas várias arestas se os ângulos forem muito próximos.

Na figura A.1 podemos verificar que o número total de segmentos aumenta aproximadamente 1000% entre intervalos, nos segmentos não duplicados esta tendência não se verifica. O aumento na percentagem de segmentos não duplicados, na passagem do valor 1 para 0.1 é na ordem dos 204%, mas nas passagens seguintes são de 72% e 31%, não havendo assim um aumento tão acentuado, conforme se pode verificar.

Para efetuar uma avaliação visual, dos segmentos não repetidos, foram criadas imagens com os segmentos de cada intervalo, 1, 0.1, 0.01 e 0.001, que correspondem respectivamente às figuras A.2(a), A.3(a), A.4(a) e A.4(b). De seguida foi produzida uma imagem, com a diferença de duas imagens consecutivas.

A figura A.2(c), representa as diferenças entre a imagem criada com o intervalo 1 e 0.1. Aqui podemos verificar que entre os dois intervalos existe informação acrescentada pelo intervalo 0.1, teremos de ter em atenção que os 4926 segmentos que a figura A.2(b) fornece, representam somente 5756 píxeis em relação à figura A.2(a).

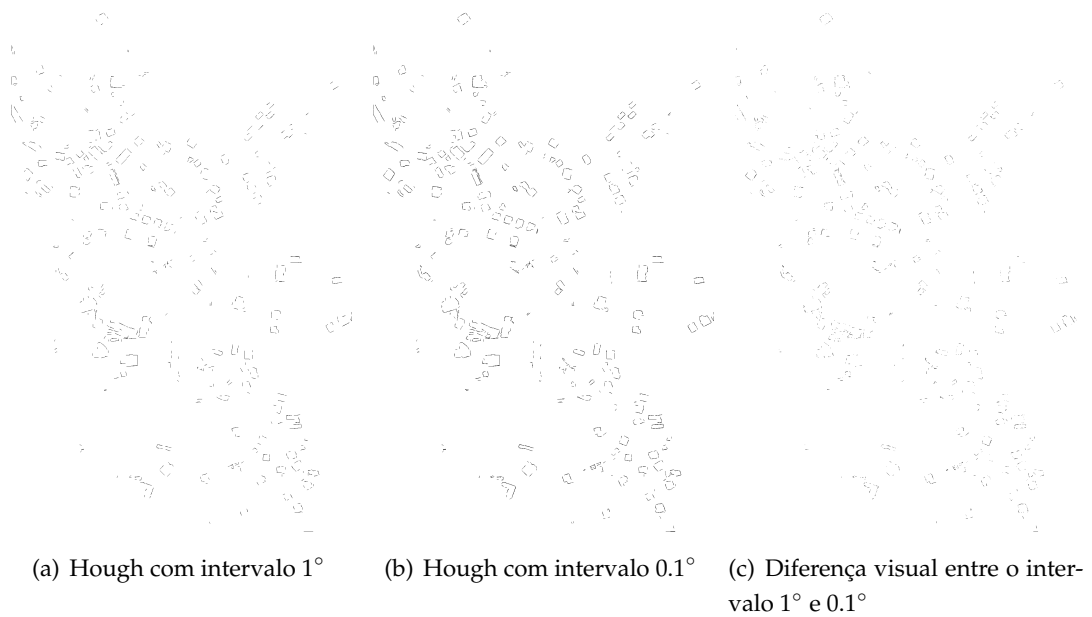


Figura A.2: Visualização das diferenças entre os intervalos 1° e 0.1°

A figura A.3(c) representa as diferenças entre as imagens com intervalo 0.1 e 0.01, sendo esta composta por 3071 píxeis. No entanto a figura A.2(b) tem mais 5064 segmentos que a figura A.2(a).

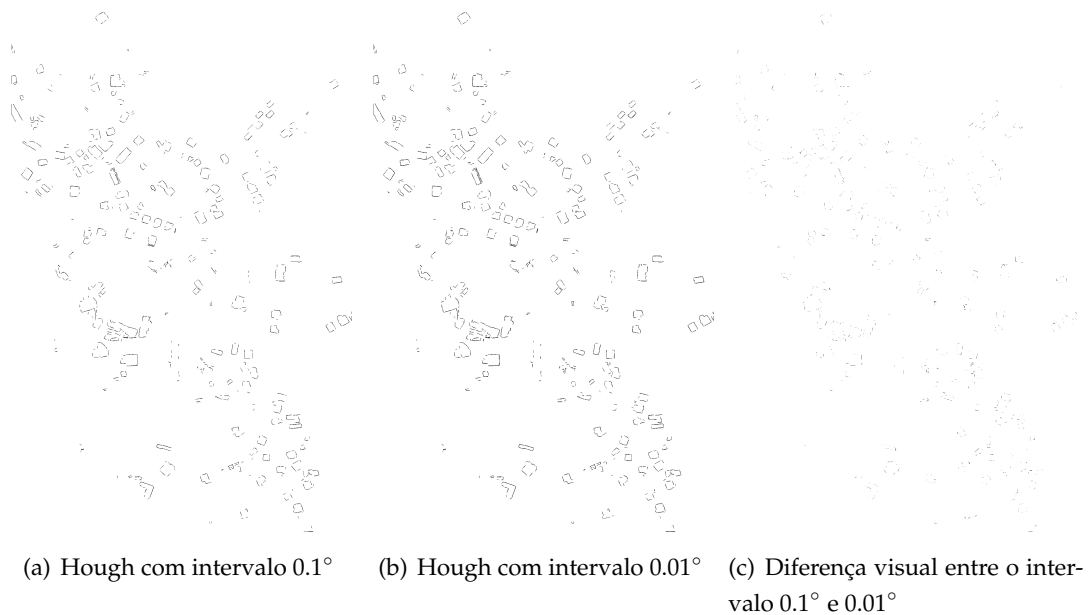


Figura A.3: Visualização das diferenças entre os intervalos 0.1° e 0.01°

A figura A.4(c) representa as diferenças entre as imagens com intervalo 0.01 e 0.001, sendo esta composta por 1383 píxeis. No entanto a figura A.3(b) tem mais 3972 segmentos que a figura A.2(b). Nesta figura praticamente não se consegue visualizar informação, devido às figuras A.3(a) e A.3(b) serem bastante semelhantes.

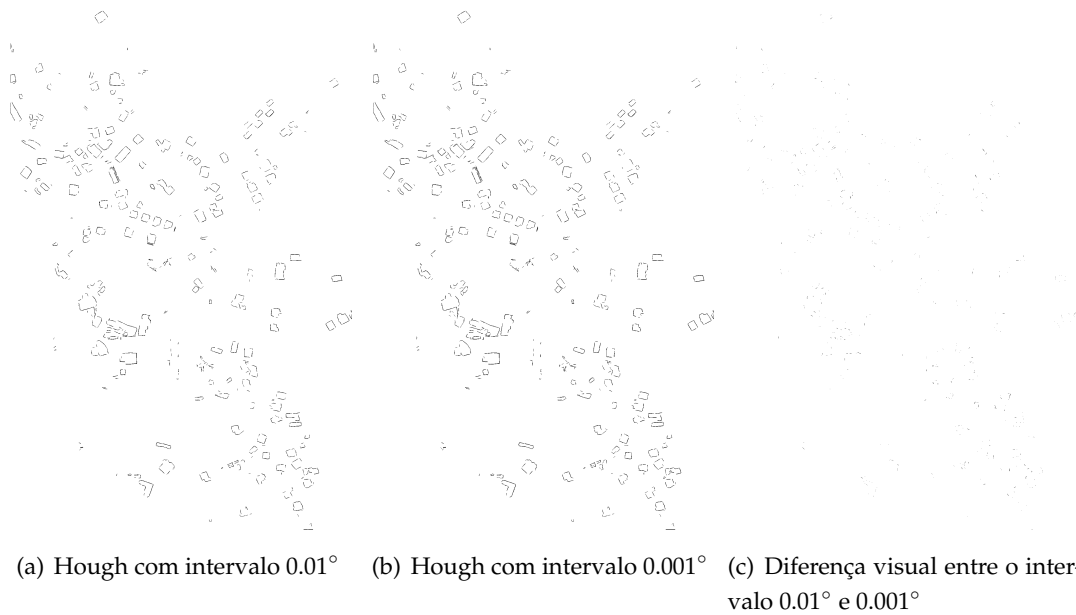


Figura A.4: Visualização das diferenças entre os intervalos 0.01° e 0.001°

Pretendemos com esta avaliação demonstrar que quanto menor for o intervalo mais informação é recolhida, mas o tempo que é despendido não compensa o acréscimo fornecido por esta informação. Assim consideramos que o valor mais apropriado a colocar neste parâmetro é 0.01 no incremento do intervalo para percorrer o ciclo entre 0° e 90° .



Análise equações de colineariedade

Neste anexo pretendemos demonstrar, que a determinação da coordenada 3D através das equações de colinearidade, esta sujeita a alterações com a variação da coordenada de destino.

Ao fixar as coordenadas origem e efetuar a variação das coordenada destino em 0.1, iremos obter alterações significativas na coordenada terreno obtida. Para concretizar utilizamos uma coordenada origem fixa com o valor $(x, y) = (1083.44, 1921.00)$ e efetuamos a variação da coordenada destino com o valor $(x, y) = (1059, 1934)$.

Na tabela B.1, podemos observar que à medida que a coordenada destino vai sendo incrementada, as coordenada 3d sofrem alterações. Esta alteração é mais significativa no valor de z

Tabela B.1: Classificação dos polígonos

Coordenada destino	Coordenada 3D	Diferença x	Diferença y	Diferença z
(1059.00, 1934.00)	(180637.56, 509019.61, 490.59)			
(1059.10, 1934.10)	(180637.58, 509019.64, 490.39)	-0.01	-0.03	0.20
(1059.20, 1934.20)	(180637.59, 509019.67, 490.19)	-0.03	-0.06	0.40
(1059.30, 1934.30)	(180637.61, 509019.70, 489.99)	-0.04	-0.09	0.59
(1059.40, 1934.40)	(180637.62, 509019.73, 489.79)	-0.05	-0.12	0.79
(1059.50, 1934.50)	(180637.63, 509019.76, 489.60)	-0.07	-0.15	0.99
(1059.60, 1934.60)	(180637.65, 509019.79, 489.40)	-0.08	-0.18	1.19
(1059.70, 1934.70)	(180637.66, 509019.82, 489.20)	-0.10	-0.21	1.39
(1059.80, 1934.80)	(180637.67, 509019.85, 489.00)	-0.11	-0.24	1.59
(1059.90, 1934.90)	(180637.69, 509019.88, 488.80)	-0.12	-0.27	1.78

Enquanto a variação de x e y não provocam uma diferença substancial na comparação

com os valores restituídos, em z a diferença já se encontra na ordem dos metros.



Imagens da análise por fases

Neste anexo encontram-se figuras que representam aspetos que não foi possível introduzir no decorrer do texto devido a ocuparem um espaço considerável. Estas foram construídas a partir do processamento da fotografia *seta01534.tif* sobre o espaço de cores HSV com cinco cores.



Figura C.1: Quantização do processamento com HVS 5

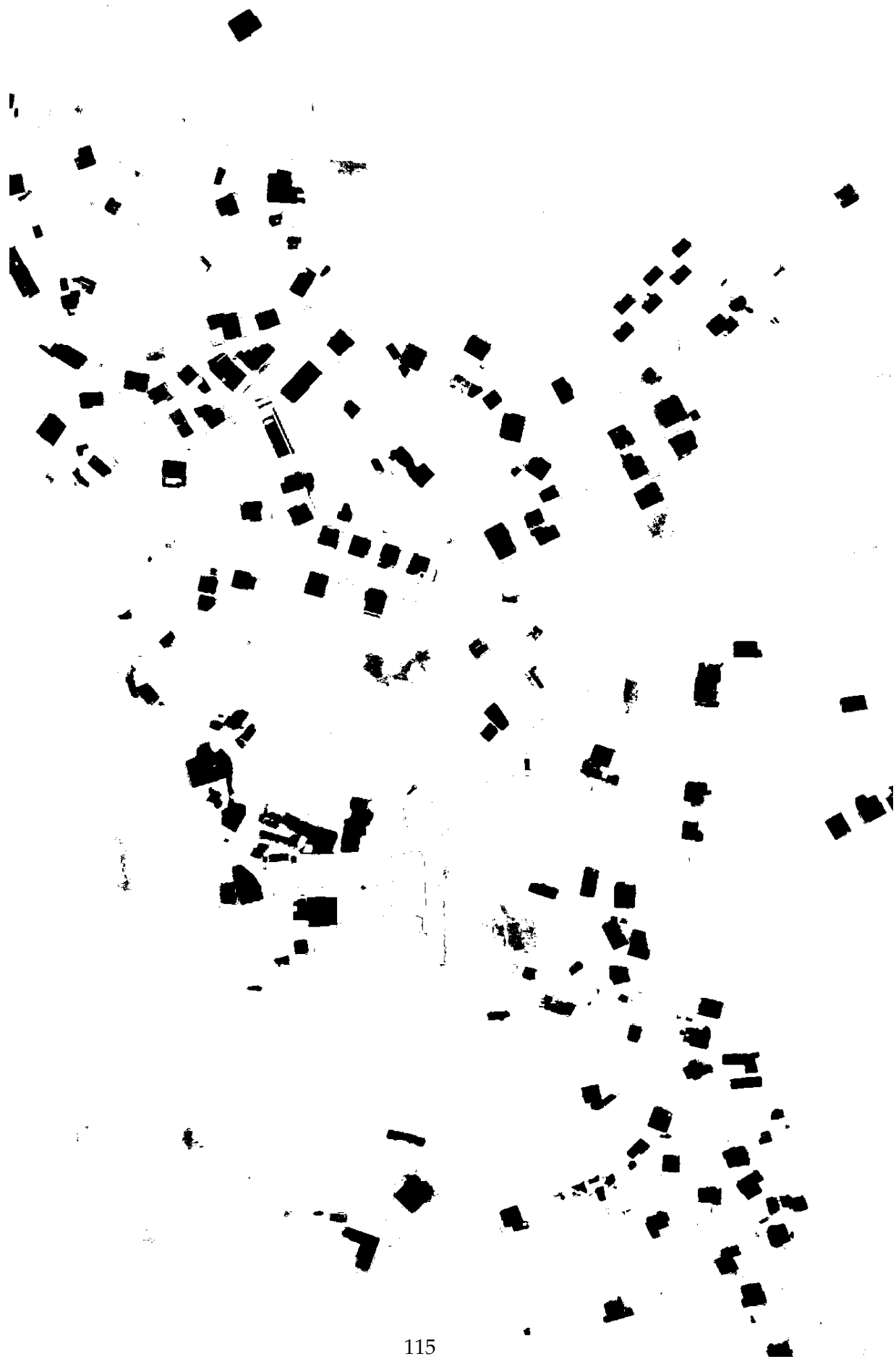




Figura C.3: Contornos do processamento com HVS 5



Figura C.4: Polígonos do processamento com HVS 5



Figura C.5: Triangulação de Delaunay utilizada na determinação de coordenadas com HVS 5



Imagens da taxa de aceitação

Neste anexo encontram-se figuras dos polígonos utilizados para a elaboração da taxa de aceitação, que não foi possível introduzir no decorrer do texto devido a ocuparem um espaço considerável. Estas foram construídas a partir do processamento da fotografia *seta01534.tif* sobre o espaço de cores HSV com cinco cores.

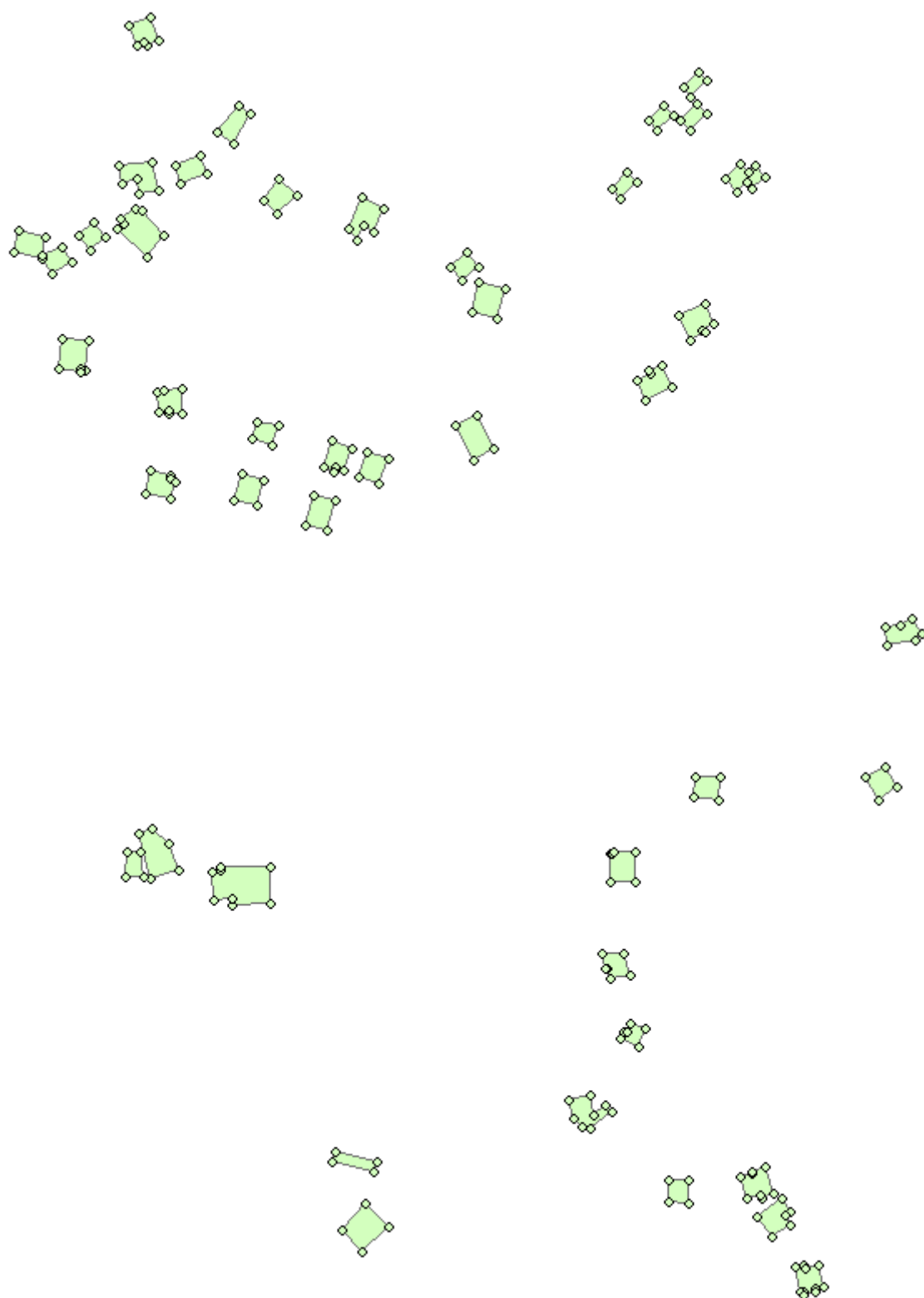


Figura D.1: Polígonos e respetivos vértices restituídos

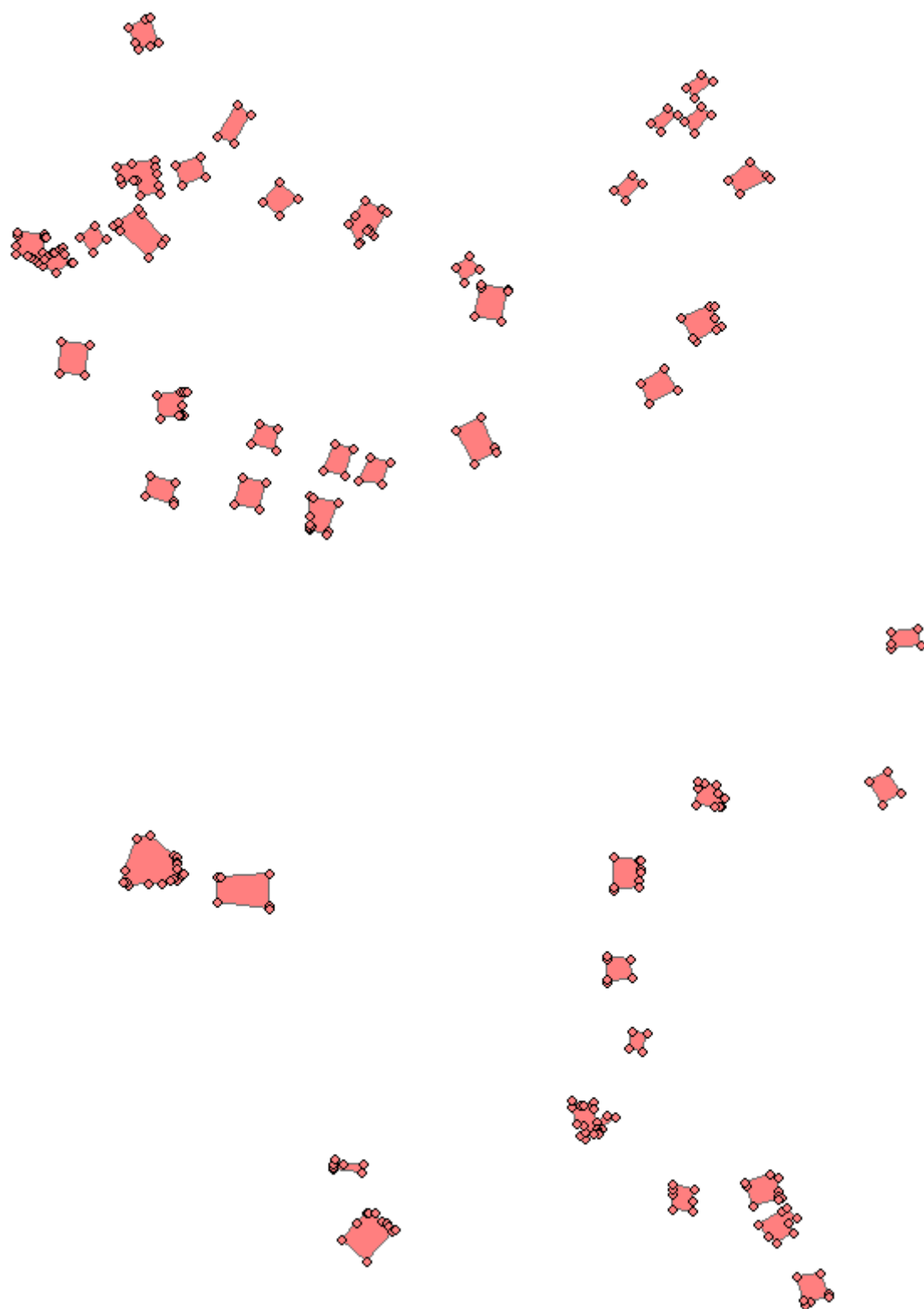


Figura D.2: Polígonos e respetivos vértices gerados

Bibliografia

- [22102] STANAG 2215. Evaluation of land maps, aeronautical charts and digital topographic data, 2002.
- [AGA08] V Arévalo, J González, e G Ambrosio. Shadow detection in colour high resolution satellite images. *International Journal of Remote Sensing*, 29(7):1945–1963, Abril 2008.
- [Bal13] Sónia Baldaia. Estudo do potencial das imagens estereoscópicas WorldView 2 na extração de informação cartográfica para a Carta Militar 1:25 000. *Boletim IGeoE*, pág. 70–82, 2013.
- [BCKO08] Mark de Berg, Otfried Cheong, Marc van Kreveld, e Mark Overmars. *Computational Geometry: Algorithms and Applications*. Springer-Verlag TELOS, Santa Clara, CA, USA, 3rd ed. edition, 2008.
- [BETVG08] Herbert Bay, Andreas Ess, Tinne Tuytelaars, e Luc Van Gool. Speeded-up robust features (surf). *Comput. Vis. Image Underst.*, 110(3):346–359, Junho 2008.
- [BK08a] Dr. Gary Rost Bradski e Adrian Kaehler. *Learning OpenCV, 1st Edition*. O’Reilly Media, Inc., first edition, 2008.
- [BK08b] G Bradski e A Kaehler. *Learning OpenCV*. pág. 70–82, 2008.
- [C62] H.P.V. C. Method and means for recognizing complex patterns, Dezembro 18 1962. US Patent 3,069,654.
- [Can86] J Canny. A computational approach to edge detection. *IEEE transactions on pattern analysis and machine intelligence*, 8(6):679–98, Junho 1986.
- [Car99] Luiz A. Cardoso. Computer-aided recognition of man-made structures in aerial photographs. Tese de Mestrado, Naval Postgraduate School Monterey, CA 93943-5000, 1999.

- [CGZ08] Hamilton Y. Chong, Steven J. Gortler, e Todd Zickler. A perception-based color space for illumination-invariant image processing. *ACM Trans. Graph.*, 27(3):61:1–61:7, Agosto 2008.
- [CLF10] Michael Calonder, Vincent Lepetit, e Pascal Fua. Brief: Binary robust independent elementary features, 2010.
- [CM99] D Comaniciu e Peter Meer. Mean shift analysis and applications. ... *Vision, 1999. The Proceedings of the ...*, (2):70–82, 1999.
- [Cyg07] B. Cyganek. *An Introduction to 3D Computer Vision Techniques and Algorithms*. John Wiley & Sons, 2007.
- [Dar05] Paul M Dare. Shadow Analysis in High Resolution Satellite Imagery of Urban Areas. *Photogrammetric Engineering & Remote Sensing*, 71(2):169–177, Fevereiro 2005.
- [DT05] Navneet Dalal e Bill Triggs. Histograms of oriented gradients for human detection. In Cordelia Schmid, Stefano Soatto, e Carlo Tomasi, editores, *International Conference on Computer Vision & Pattern Recognition*, volume 2, pág. 886–893, INRIA Rhône-Alpes, ZIRST-655, av. de l’Europe, Montbonnot-38334, June 2005.
- [dT13] Direção-Geral do Território. *Normas técnicas de produção e reprodução*. 2013.
- [EWa] Weisstein Eric W. Cevian. From MathWorld—A Wolfram Web Resource. Last visited on 14/8/2014.
- [EWb] Weisstein Eric W. Euler angles. From MathWorld—A Wolfram Web Resource. Last visited on 14/8/2014.
- [FB81] M Fischler e R Bolles. Random Sample Consensus: A Paradigm for Model Fitting with. *Communications of the ACM*, 1981.
- [FH06] K. Fukunaga e L. Hostetler. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Trans. Inf. Theor.*, 21(1):32–40, Setembro 2006.
- [FPAW14] R Fisher, S Perkins, e E Wolfart A Walker. Affine transformation, 2014. [Online; accessed 14-December-2014].
- [fris14] Centre for remote imaging sensing e processing. Interpreting optical remote sensing images, 2014. [Online; accessed 27-Novembre-2014].
- [GHS01] M Gerke, C Heipke, e B.-M. Straub. Building extraction from aerial imagery using a generic scene model and invariant geometric moments. *IEEE/IS-PRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas (Cat. No.01EX482)*, pág. 85–89, 2001.

- [GJMR12] Rafael Grompone von Gioi, Jérémie Jakubowicz, Jean-Michel Morel, e Gregory Randall. LSD: a Line Segment Detector. *Image Processing On Line*, 2:35–55, Março 2012.
- [Gon06] Gil Rito Gonçalves. *Elementos de Fotogrametria Analítica*. Coimbra, 2006.
- [GS99] Theo Gevers e AWM Smeulders. Color-based object recognition. *Pattern recognition*, 32(February 1998):453–464, 1999.
- [Han05] Jiawei Han. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005.
- [HB05] Heiko Hirschm e P O Box. Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. 2005.
- [HN88] a Huertas e R Nevatia. Detecting buildings in aerial images. *Computer Vision, Graphics, and Image Processing*, 41(2):131–152, Fevereiro 1988.
- [HZ03] R Hartley e A Zisserman. Multiple View Geometry in Computer Vision. pág. 237–247, 2003.
- [IGe08] IGeoE. *Normas Técnicas de apoio à aquisição de dados*. 2008. [AAL015 e LBH015].
- [IHKM12] NA Ibraheem, MM Hasan, R Z Khan, e P K Misha. Understanding Color Models: A Review. *ARPJ Journal of Science and Technology*, 2(3):265–275, 2012.
- [IM89] R B Irvin e D M McKeown. Methods for exploiting the relationship between buildings and their shadows in aerial imagery. *OE/LASE’89, ...*, 1989.
- [LMS06] B. Leibe, K. Mikolajczyk, e B. Schiele. Efficient Clustering and Matching for Object Class Recognition. *Proceedings of the British Machine Vision Conference 2006*, pág. 81.1–81.10, 2006.
- [LN98] Chungan Lin e Ramakant Nevatia. Building Detection and Description from a Single Intensity Image. *Computer Vision and Image Understanding*, 72(2):101–121, Novembro 1998.
- [Low04] DG Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, Novembro 2004.
- [May99] Helmut Mayer. Automatic Object Extraction from Aerial Imagery—A Survey Focusing on Buildings. *Computer Vision and Image Understanding*, 74(2):138–149, Maio 1999.
- [MGK98] J Matas, C Galambos, e J Kittler. Progressive Probabilistic Hough Transform. *BMVC*, 1998.

- [ML09] Marius Muja e DG Lowe. Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. *VISAPP (1)*, 2009.
- [MLB⁺03] MA Maloof, P Langley, TO Binford, R Nevatia, e S Sage. Improved rooftop detection in aerial images with machine learning. *Machine Learning*, pág. 1–44, 2003.
- [MLSB97] Marcus Maloof, Pat Langley, Stephanie Sage, e Thomas O. Binford. Learning to detect rooftops in aerial images, 1997.
- [MS05] Krystian Mikolajczyk e Cordelia Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 27(10):1615–1630, 2005.
- [NB80] Ramakant Nevatia e K Ramesh Babu. Linear feature extraction and description. *Computer Graphics and Image Processing*, 13(3):257 – 269, 1980.
- [NLH97] R Nevatia, C Lin, e A Huertas. A system for building detection from aerial images. . . .-made objects from aerial and space images . . . , 1997.
- [Ots79] Nobuyuki Otsu. A Threshold Selection Method from Gray-level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1):62–66, 1979.
- [Pes06] António Pestana. *Elementos de Topografia - Volume 1*. 2006.
- [PGB03] Patrick Pérez, Michel Gangnet, e Andrew Blake. Poisson image editing. *ACM SIGGRAPH 2003 Papers on - SIGGRAPH '03*, pág. 313, 2003.
- [Poy96] Charles A. Poynton. *A technical introduction to digital video*. Wiley, 1996.
- [PV00] Konstantinos N. Plataniotis e Anastasios N. Venetsanopoulos. *Color Image Processing and Applications*. Springer-Verlag New York, Inc., New York, NY, USA, 2000.
- [Red07] Paula Redweik. *Fotogrametria aérea*. FCUL, 2007.
- [Red10] Paula Redweik. *Fotogrametria analítica*. FCUL, 2010.
- [SA85] S Suzuki e K Abe. Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing*, 46:32–46, 1985.
- [SAH08] C. Silpa-Anan e R. Hartley. Optimised kd-trees for fast image descriptor matching. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pág. 1–8, June 2008.
- [Sal06] David Salomon. *Data Compression: The Complete Reference*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

- [San07] Carla Santos. *Estatística Descritiva - Manual de Auto-aprendizagem*. Edições Sílabo, 2007.
- [Shi09] Jianbo Shi. Building recognition from aerial images combining segmentation and shadow. *2009 IEEE International Conference on Intelligent Computing and Intelligent Systems*, pág. 578–582, Novembro 2009.
- [ST10] DK San e M Turker. Building extraction from high resolution satellite images using Hough transform. *International Archives of the . . .*, XXXVIII(2001):1063–1068, 2010.
- [SU08] B. Sirmacek e C. Unsalan. Building detection from aerial images using invariant color features and shadow information. In *Computer and Information Sciences, 2008. ISCIS '08. 23rd International Symposium on*, pág. 1–5, Oct 2008.
- [SWD05] Gaurav Sharma, Wencheng Wu, e Edul N. Dalal. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application*, 30(1):21–30, Fevereiro 2005.
- [TCT⁺13] Gabriel Takacs, Vijay Chandrasekhar, Sam Tsai, David Chen, Radek Grzeszczuk, e Bernd Girod. Rotation-invariant fast features for large-scale recognition and real-time tracking. *Image Commun.*, 28(4):334–344, Abril 2013.
- [Wor] Worqx. Color systems. From worqx.com. Last visited on 1/9/2014.